Subjectivity and Ethics

Paul Ricoeur and the Question of Naturalizing Personhood

Rosfort, René

Publication date:
2008

Document version
Publisher's PDF, also known as Version of record

DET TEOLOGISKE FAKULTET
KØBENHAVNS UNIVERSITET

RENÉ ROSFORT

Subjectivity and Ethics

Ricoeur and the Question of Naturalizing Personhood

ISBN 978-87-91838-09-05

RENÉ ROSFORT

Subjectivity and Ethics

Ricoeur and the Qustion of Naturalizing Personhood

Publikationer fra Det Teologiske Fakultet

Subjectivity and Ethics

Ricoeur and the Question of Naturalizing Personhood

Ph.D.-Thesis by René Rosfort

Funded by Copenhagen University's Research Priority Area
Body and Mind
Submitted for defence at the Faculty of Theology
University of Copenhagen
1 February 2008

Subjectivity and Ethics
Ricoeur and the Question of Naturalizing Personhood
Publikationer fra Det Teologiske Fakultet 2

ISBN: 978-87-91838-09-05 (trykt)
ISBN: 978-87-91838-98-9 (pdf)
Københavns Universitet 2008

Udgivet af
Det Teologiske Fakultet
Københavns Universitet
Købmagergade 44-46
1150 København K

Contents

Part Two
Feeling Ethical

Chapter One: Human Affectivity

Chapter Two: Ethical Experience and Ontology

Part Three
The Brain and Human Values

Chapter One: Neuroethics and Subjectivity

Chapter Two: Personhood in Nature and Culture

Introduction

Subjectivity and ethics are both difficult topics to reconcile with a empirical scientific outlook on the world and the creatures living in it. They are somewhat impalpable subject-matters in a discourse that enhances and applauds local, observable, and experimental-proved research and which is, to put it mildly, skeptical of attempts to explain human behavior in the light of first-person perspective or general principles of morality. Theories of ethics and subjectivity have traditionally included such non-observable perspectives and have, indeed, regarded them as fundamental for ethical debate or more general investigations into human nature. We cannot observe a rational principle or a moral sentiment in the same manner as we observe the movement of the cochlea of the inner ear in response to sound waves; nonetheless, principles and sentiments seem to have some kind of influence on how we, as human beings, respond to our environment. Empirical sciences and theoretical investigations into subjectivity and ethics have therefore always faced (and continue to face) the problem of integrating their different methods, analyses, and results.

This problem has existed throughout the history of western thought, stressed with the scientific revolution in the seventeenth century and reemerged as a dominant issue over the last forty years. Confronted with the prodigious growth of technological developments and scientific revelations in the twentieth century, philosophers and empirical scientists have developed various attitudes toward the problem of integration, from mutual dialogue to reduction or even complete elimination of the position of the other. Although mutual dialogue might seem the most sensible answer, the debate has often opted for the radical pole of complete elimination. To get a feeling of this, we can look at what two very influential thinkers had to say about the problems of integration in the last century.

In 1929 the philosopher Martin Heidegger launches a polemic attack on the scientific enthusiasm blooming in his time: 'No time has known so much and such a variety about mankind as is the case today […] But also, no time has known less about what man is than today. In no other time has man become as questionable as in ours' (Heidegger 1998: 143/209). The empirical sciences can perhaps inform us with all sorts of interesting things about mankind, but it can never solve the general problems of being human.

Almost forty years later, the paleontologist George G. Simpson wrote, in a no less polemical tone, against the metaphysically (i.e. not based on scientific methods and the theory of evolution) dreaming philosopher: 'The question "What is Man" is probably the most profound that can be asked by man. […] The point I want to make now is that all attempts to answer that question before

1859 [the publication of Darwin's *The Origin of Species*] are worthless and that we will be better off if we ignore them completely' (Simpson 1966, 472).

Although Simpson later in the article tends to give some kind of credit to '[t]he other, older approaches through metaphysics, theology, art, and other nonbiological, nonscientific fields', he persists in considering them 'merely fictional fantasies or falsities' (op.cit., 473) if they do not accept that human nature is biological organism.

This conviction of an abiding necessity to ground all investigations of human nature in biology has become even more prevailing today. The explosive development in neuroscience has infused new hope into the old idea of an entirely naturalized conception of man. New techniques, intriguing results, and consequently large fundings have paved the way for the so-called neuroscientific revolution that in the last thirty years or so has prophesized the arrival of a new, more sincere, conception of human nature (Edelman 2003: 5520; Metzinger 2005: 54; Pinker 1999: 563; Pinker 2003: xi; Gazzaniga 2006: xviii). Together with the growth of cognitive science, socio-biology, and lately evolutionary psychology, neuroscience is expanding its field of interest (and presumed expertise) from neurophysiological investigations into specific neuronal networks and brain areas to more general theories about human nature.

So, this said, what are the problems with the attempt to naturalize human nature? The problems are spread all over the philosophical spectrum, but the main debate has always taken place in philosophy of mind. With our growing empirical knowledge of the structure and dynamics of the brain, we cannot just dispose of that knowledge when we think about the mind. The mind-body problem has, once again, become the battlefield of empirical sciences and philosophy – this time in the shape of a mind-brain problem. The debate has been, and still is, a very tense and emotionally-loaded affaire, since it is our very conception of human nature that is at stake. Often, philosophy is ridiculed as mere daydreaming and science accused of short-sighted superficiality. But the divide exists vividly within the disciplines themselves. There are many neuroscientists who caution against the holistic tendencies within their own discipline, whereas some prominent philosophers emphasize the necessity of doing philosophy on the terms of the empirical sciences (e.g. Dennett, Metzinger, Paul and Patrica Churchland).

As mentioned, one major point of conflict is how to integrate the scientific methods with the supposedly non-scientific ones regarding, among many others, the exploration of subjective experiences and feelings such as moral principles, ideas of happiness and good education, sadness, despair, love, redness, shame, joy, anger, in short, our first-person (or perhaps it is better to say

personal) perspective on ourselves and the world. This is the question of what philosophers traditionally have named "qualia", i.e., the distinctive subjective character of experience. Nobody would deny that "qualia" exists, but many question their ontological status. They have been thought of as mere epiphenomena and folk fantasies, or as one of their most acerbate antagonists, Daniel Dennett, puts it: 'Qualitative properties that are intrinsically conscious are a myth, an artefact of misguided theorizing' (Dennett 1991: 40). Now, such a denial of these qualitative features of experience is problematic, because it often leads to the denial of the subjects that have the experiences. And this denial has critical implications for the way we understand and deal with the nature of human beings. On the other side, we have philosophers (together with theologians and other branches of the humanities) that question the legitimacy of the project of naturalization. They argue that we cannot accept explanations of human nature that does not account for the specific nature of subjective experience but merely reduce them to disturbing by-products or epiphenomenal illusions of a more true and solid physical world. The subjective features of experience possess a legitimate claim on investigation, because they reveal the nature of a self, which is an indispensable part of being human. As the philosopher Dan Zahavi writes in answer to Dennett's harsh dismissal of subjectivity: 'a science of consciousness should draw on both the first-, second- and the third-person point of view, just like all of us do when we engage in the everyday practice of understanding ourselves and others' (Zahavi 2007: 38; see also Nagel 1986: 15; Searle 2000: 42-3; Taylor 1977b: 37).

This leads me to the issue that I shall treat in the following chapters, namely the relation between subjectivity and ethics with regard to the problems of naturalization.

Subjectivity, Ethics, and Naturalization

Subjectivity and ethics both have a long history, and they have often crossed each other up through the centuries. The central question that forms the background for the analyses in the present work is the following: can we understand subjectivity without dealing with ethics, and vice versa? I shall argue that we cannot. This, however, is by no means an intuitive fact; on the contrary, the battle has been going on since Socrates first attacked the sophists and their immoral self-contradictions regarding the subjectivity of good action. Aristotle went for a subtle middle-position that emphasized the importance of human nature in dealing with ethics, and yet he still worked with a certain objectivity regarding human capacities and behavior. Hume and Adam Smith argued for the subjective nature of morality, whereas Kant and Moore showed that good and bad are somehow

independent of our subjective experience of emotions, inclinations, and reality. And further, the contemporary debate shows that the question is still a relevant one to ask. Now, the question can be asked from at least two perspectives, namely from ethical theory and from theories of subjectivity.

With regard to ethical theory, few, and perhaps no, moral philosophers would question the importance of human nature in ethical debate; the real issue what we understand by human nature. How much of what it is to be a human being depends on subjective experience? Does our subjective coloring of the world have anything to do with how we should behave? Are moral principles somehow detached from the baser workings of desire and feeling? These questions might be answered with the claim that only firm principles are valid; or, for example, that morality is grounded in a transcendent god who has revealed the infallible rules for human conduct, which have nothing to do with the wretched and sinful nature of the individual subjects; or that human experience of good and bad is always tainted with selfish and futile motives that do not promise anything but total chaos, so morality cannot but be grounded in something raised above subjective experience; or that the morally good transcends, in some particular way, our subjective experience of certain qualities, desires and inclinations and is completely different from other non-moral properties in the world. Others, however, retain that subjective experience has a lot to say about the nature of moral norms and values; in fact, they believe that we cannot say anything about ethics without an attentive analysis of our subjective experience of the world, the other subjects and ourselves. Subjective experience matters because it is part of who we are and what we do, and dismissing it as irrelevant to ethics is, as the moral philosopher Simon Blackburn writes, the same as to discredit 'any theory that seeks to explain our moral capacities in terms of contingent and potentially variable facets of human nature: language, culture, upbringing, acquired "second nature", and so on' (Blackburn 2006: 145). I shall defend the latter position and try to argue that a thorough investigation into the nature of subjectivity will benefit the contemporary metaethical debate concerning the relation between human nature and morality.

From the point of view of subjectivity theory, ethical questions have had a quite ambiguous status. Subjectivity became an explicit theme with the coining of phenomenology. Of course, subjectivity has been discussed since antiquity (especially with the sophists and the skeptics), and in modern times Locke and Leibniz took the notion under serious consideration, whereas Hume, Kant and Kierkegaard paved the way for a systematic treatment. Nevertheless, it was with the prodigious writing of the philosopher Edmund Husserl in the beginning of the last century that the notion became a distinct field of investigation. Husserl constituted phenomenology as a clarification of the

structures and dynamics of subjective experience in relation to human perception, knowledge, and action. His works influenced generations of philosophers afterwards, which among the most influential include Scheler, Heidegger, Merleau-Ponty, Ricoeur, Sartre, Lévinas, and Derrida. Today, phenomenology still defends the importance of the notion of subjectivity and first-person perspective in disciplines such as philosophy of mind, sociology, psychiatry, and neuroscience (Zahavi 2005; Gallagher and Zahavi 2008). Ethics, however, has not been a frequent theme in phenomenology. Apart from the works of Scheler, Ricoeur, and Lévinas, subjectivity has been investigated without explicitly considering the subject's ethical dispositions. The focus has been on perception, knowledge and theory of action, and even though the question of the other has been recent topic in the literature, the problem of value and morality has never really gained a proper articulation inside the phenomenological debate. When contemporary phenomenologists discuss ethics it is normally in form of kind of 'proto-ethics' that describes the subject's openness and structural relationship with (and infinite responsibility to) the other subject (Overgaard 2007: 143ff). The systematic treatment of the subject's values and norms tends to be considered as an extra dimension that can be left out of the structural analyses of subjectivity. On the following pages, I will argue that it cannot. We might suspend the explicit normative aspect of ethical theory, that is, which set of norms are the most adequate, just or right, in a certain society, a given situation, or with respect to different cultures etc., but we cannot dismiss (meta)ethical problems such as how the subject experiences, responds to, and act upon values. Why does perception of a supposedly non-normative world become faceted with normative aspects? How are values generated in subjective experience?

The central parts of this work are dedicated to the relation between subjectivity and ethics. It is not about which kind of ethics such an emphasis on the notion of subjectivity would eventually lead to. The analyses of subjectivity will be held on a descriptive level, since they aim at showing that ethical key concepts such as values, normativity, the other, and practices are part of the factual existence of the human subject. The analyses focus on such key-concepts and do not deal explicitly with ethics. It might be said to be a metaethical investigation of the subjective nature of ethical concepts and an investigation into the ethical dimension of subjectivity. There will be no formal distinction between these two aspects of the dissertation, since both are clarified simultaneously in the progress of the analyses. In fact, it is the central aim of the dissertation to show that these two aspects cannot be dealt with in isolation, but only through an attentive analysis of the relation that bind them together.

Structures of the Analyses

The first part is dedicated to a reformulation of Paul Ricoeur's theory of subjectivity, since his work is an important contribution to the discussion of subjectivity and ethics. He argues for an inseparable relation between the two and formulates arguments that with benefit can be used in the contemporary debate. The second part develops some aspects Ricoeur's theory with further analyses human affectivity and ethical experience. Ricoeur's analyses and arguments are used as the framework in which these developments are done. I bring in the neurophysiological dimension of affectivity, contemporary theories of emotions, and contemporary metaethical analyses of the relation between feelings, values, and practices. I argue for a naturalistic conception of subjectivity that understands ethics as a natural part of the subjective being of the human person. The third part deals with the attempt to naturalize human personhood and ethics. I hope clarify my own naturalistic position by criticizing both the new born discipline neuroethics and a recent argument for moral relativism. This part is shorter than the other two, but although I only approach the question of naturalization directly in the last thirty some pages, the question is present in all the analysis, if only in the sense of preparing an answer. In particular the second part paves the way for a treatment of the question about naturalization in regard to subjectivity and ethics, since it looks upon neurophysiological dimension of affectivity and values in relation to human practices.

Thus, to summarize and anticipate the problems that will be dealt with in the following chapters:

- **A Reformulation of Ricoeur's Theory of Subjectivity**
  1. The Schematic Notion of Subjectivity
  2. The Redressed Notion of Subjectivity
- **Feeling Ethical: The Relation between Subjectivity and Ethics**
  1. Human Affectivity
  2. Ethical Experience
- **Naturalized Personhood: The Brain and Human Values**
  1. Neuroethics: Subjectivity and Nature
  2. Personhood in Nature and Culture

I thank from the heart my colleges at Center for Subjectivity Research, my supervisors Arne Grøn, Hans Hultborn, and further Giovanni Stanghellini, Roberta Lanfredini, il Dipartimento di Filosofia a Firenze, Jacob Simonsen, and Iben Damgaard for inestimably competent help and warm support.

Part one

Ricoeur's Theory of Subjectivity

The central purpose of this part is to clarify Paul Ricoeur's theory of subjectivity. The notion of the subject is of primary importance in nearly all of Ricoeur's writings, and I believe that the basic structures and constitution of subjectivity can be analyzed systematically without necessarily covering his philosophy in general. In fact, I believe that a theoretic reformulation of the notion of subjectivity might be a help if one wants to approach his works as a development of single basic ideas. An enormous, life-long, quantity of writings in phenomenology, semeiotics, hermeneutics, social sciences, historical theory, theology, and ethics can make it difficult to find a theoretical core in his work. Some might disagree, but I think that the notion of subjectivity is a good place to start in such an attempt to emphasize importance of his theoretical groundwork. The following parts will show how his theory of subjectivity can be used with good profit in debates where one rarely finds his name.

Contrary to Ricoeur's own method, I shall make an effort to reformulate his theory of subjectivity without taking into account the vast number of dialog partners that Ricoeur himself uses as a vehicle for his thinking. Furthermore, the interpretation will not deal explicitly with the chronological development of the single concepts nor of his philosophy in general. Apart from the fact that the interpretation is divided in two chapters that deal with two different periods of his production (and therefore the analyses are done with reference to the works from the respective period), his works will be used, for the most part, without explaining the argumentative transitions between the different works. This choice is not only due to the limited space available but because I believe that his theory has its own theoretical grounding which deserves to be clarified. Besides that, several valuable studies have previously dealt with his so-called *voie longue* seeking to pin out the implications of his various dialogues with past and contemporary thinkers[1]. This choice is not, in any way, a disproval of neither the hermeneutical nor the comparative method as an approach to Ricoeur's philosophy, but the consequence of a strictly systematic interest. I shall use his theory as a theoretical framework for a better understanding of the relation between subjectivity and ethics and in a concluding discussion of the naturalization of human personhood. Once again, the following interpretation is neither concerned with Ricoeur's philosophy in general nor with his much-discussed standpoint in contemporary philosophy, literary-theory, social sciences or theology.

[1] Among many others: Greisch (2000, 2001), Mongin (1996), Jervolino (1993, 1995), Brezzi (2006).

It is an attempt to clarify the basic structures in the constitution of the subject as they are conceived by Ricoeur and developed throughout his writings in the last fifty years. The reformulation will necessarily appear unsatisfactory and perhaps even poor in the eyes of people that know Ricoeur's work better than I do, but they must keep in mind the general themes of the work, while reading the following pages. I try to use Ricoeur's analyses and ideas in a context that is different from his own methods and conceptual background.

I have therefore confined myself to a limited selections among his works. I deal explicitly with six major works: 1) *Freedom and Nature*, 2) *Fallible Man*, 3) *Freud and Philosophy: An Essay on Interpretation*, 4) *Time and Narrative vol. 3*, 5) *Oneself as Another*, and 6) *The Course of Recognition*[2]. These works are chosen according to two criteria. First, they are the most theoretical treatments of the notion of subjectivity. And secondly, they cover the earliest writings and the latest, and thereby account for the development of his theory. I shall, nevertheless, make continuously use of several of the articles and minor essays that account for a significant part of Ricœur's thinking. They often contain explanative value in the sense that they elucidate the nature of singular concepts and intricate arguments in the major works.

The analyses are divided into two chapters that deal with two significantly different approaches in his theory of the subject. This difference is main due to the hermeneutical turn that takes place after the two early works. Chapter one is an analysis of the two early writings FN and FM. Here I deal with what I have named the *stripped* notion of the subject. I use the term stripped because of the schematic analysis of the structures of the subject that takes place in these two works. In this approach, Ricoeur proceeds from experience of the world to the self-experience of the subject at three levels: from experience over action to affectivity. The second chapter examines what I have named the *redressed* notion of the subject. This is done through an analysis of FP, TN3, OSA, and CR. In these late works Ricoeur expands his notion of the subject in the sense that he relates it explicitly to the concrete, historic world and, not least, to ethics and the question of the other. Thus, he redresses the human subject with the features that he deprived it of in the early works. The arrangement of the analyses is the opposite of the former in the sense that they proceed from the subject's self-experience to the experience of the world, and that once again at three levels: from affectivity over action to ethical experience.

[2] Henceforth, I will mainly use the abbreviations listed on pp. 236-7. The quotations refer to both the English translation and the French original text. The first number indicates the English and the second the French, separated by a slanted stroke. I have modified the translation where I held that the existing translation was bad or directly misleading.

Chapter one

The Stripped Notion of Subjectivity

This chapter deals with the schematic notion of subjectivity, that is, the basic structures of the human subject, through an analysis of Ricoeur's first two major works, *Freedom and Nature* and *Fallible Man*. FN (1950) is a phenomenological exploration of the possibilities and limits the concept of volition. His studies on Husserl find their application in this work, in particular the concept of intentionality. The analyses confine themselves to an eidetic enquiry, i.e., focusing only on the human capacity to will and bracketing the question about the ontological status of the human subject, its relationship to the world and the other subjects.

Ten years later, in 1960, the next major work FM appears as a sequel to FN. In this work the brackets are removed and the ontological question comes into question. By doing so, Ricoeur seeks to place the human being in the world and look closely on how the willing subject interacts with the world and other subjects. Still, this is not a full-blown ontology, but an introduction of basic concepts for a further understanding of the human being in the world. It is in this work that Ricoeur brings Kant into his philosophy, and although he uses Kant in almost every aspect of his thinking, I believe that the Kantian concept of transcendental synthesis can be singled out as the fundamental concept for these first analyses of subjectivity.

Before turning to the explicit analysis of the stripped notion of subjectivity, it might be useful to clarify three basic methodological concepts that Ricoeur uses in these early works. Although they are introduced in the early works, these concepts remain, however, in all his writing.

Basic Methodological Concepts:

Although I am convinced that the notion of subjectivity can be reformulated without taking into account Ricoeur's various dialogue partners, one cannot avoid dealing, at least briefly, with the Kant and Husserl. They are everywhere in his writing and have influenced his thought numerous ways, but for the present interpretation their main contribution to his theory of subjectivity is found in two key concepts, namely transcendental synthesis and intentionality. The same goes for the concept of originating affirmation. However, unlike the first two, this one does not get a detailed systematic treatment in any of Ricoeur's works. Nonetheless, it is perhaps the most important of the three. It lurks, somehow, in the background of the analyses and appears under various forms from the earliest writings to the latest. It is not used nearly as explicit as the two other concepts, but confers a certain style to his way of doing philosophy (NA 378). Hence, one must be aware of how

this concept is present in his analyses and the role that it plays in his methodology. But, for now, I will start Kant and Husserl.

Ricoeur himself provides a concise description of how Kant and Husserl are combined in his philosophy: 'Husserl *did* phenomenology, but Kant *limited* and *founded* it' (KH 201/313). He uses Husserl as a methodological guide to the description of the structures of subjectivity and Kant as the speculative engine in his thinking about the ontology of subject. Thus, one might say that Ricoeur's theory of subjectivity is developed from Kant's limiting ontology and Husserl's phenomenological method. But still, his use of the two philosophers is quite unorthodox, and it is not as an interpreter of either Kant or Husserl that he has gained recognition. He coins his own notions from their forms and performs both minor and major semantic changes in the use of their concepts. I do not go into further detail about his critiques on and use of Kant and Husserl, since this has been done thoroughly in other studies, but confine myself to the basic concepts by means of which he work out his theory of the subject.

Transcendental Synthesis

The human subject is an intermediate being constituted by sensibility and reason: 'he is intermediate because he is a mixture, and a mixture because he brings about mediations' (FM 3/23). The subject and the world is understood only by turning the attention away from the experienced world and back to the subjective structures that make this experience possible in the first place. This focus on subjectivity is known as Kant's Copernican revolution and it plays a critical role in Ricoeur's analyses (FN 32/35, 471/443). In his use of the concept of transcendental synthesis Ricoeur follows straight in the footsteps of Kant.

The transcendental synthesis is the first step in understanding the subject as well as the world. Transcendental here means the basic structures of the subject that make experience of reality possible (FM 5/25). Reality is a result of the mediations performed by the subject. This is not to be misinterpreted as a kind of idealistic solipsism that reduces the reality of the world to the subjective representation of the world. On the contrary, the world is there and affects the subject through the body: 'existence of the body is the decisive fact which forces us to elucidate concrete life at the limits of intelligibility' (FN 135/129). The presence of the world is what creates a synthesis, namely the synthesis between the real existence of the world and the peculiar nature of the subject.

The subject is open to the world through its body: 'What is precisely the bewildering in the mediating role of the body is that it opens me to the world; in other words, it is the organ of an

intentional relation in which the body is not the boundary of my existence but the its correlate' (NA 307/380-81; cf. FM 20/38). Nevertheless, the subject does not coincide with its sensibility; it does not become one with the world. It is turned back upon itself in the experience of the sensible objects (FM 21/39). This is why the synthesis is *objectival*, not introspective or objective (this distinction will be explained further in the next section). It is in this being turned back upon itself, in the separation from the world, that the subject discovers itself as an intermediate being. It is not constituted by the external world alone but, other than receptivity, it possesses a spontaneity capable of representing things and ideas different from the perceived present. Spontaneity coincides with the workings of pure reason without the input from sensible data. But how do these two seemingly heterogeneous capacities come together in the unity that we know as the human subject?

The human subject is neither mere sensibility nor pure reason, but a synthesis of the two. At this point, we encounter another fundamental capacity of the human subject, namely transcendental imagination. Imagination is the third capacity that unifies the synthesis by mediating between sensibility and reason. Imagination must be homogeneous with both sensibility and reason to be able to mediate between the two. Its nature is hidden from our full comprehension and remains an enigma. Ricoeur, like Kant himself, does not pretend to decide on the nature of imagination (FM 41-2/59). He is satisfied with defining its function as generating a synthesis between the two fundamental capacities in the human subject, sensibility and reason.

The transcendental synthesis then is defined as consciousness, but a formal consciousness that is not yet self-conscious, because '[t]he consciousness philosophy speaks of in its transcendental stage constitutes its own unity only outside itself, on the object' (FM 45/63; cf. FM 18/36-7, 37/55; LS 327). I shall give a more detailed description of transcendental imagination later on. For now, suffice it to say that through the experienced object the subject understand itself as constituted by sensibility and reason (receptivity and spontaneity), but nothing more. It is an empty consciousness that has no idea of how or why it exists. However, the transcendental synthesis provides a basic scheme for the analyses of human subjectivity and is therefore the necessary first step in understanding of the nature of the subject (FM 5/25, 45-6/63).

At this point, though, the question is how we discover the subject as being a synthesis of sensibility and reason in the first place. If the subject is a being constituted neither by mere physical input received through the senses nor pure spontaneity of reason, we need a concept explaining consciousness as more than just a collection of impressions or intra-mental ideas. Ricoeur finds this explanation in Husserl's analyses of the concept of intentionality (FN 8/12, 42-3/42; SP 156-7).

Intentionality

The concept of intentionality covers a group of experiences that are all determined by their object-directedness. To be directed towards an object is a constitutive feature of human consciousness. Consciousness is always consciousness of something, a psychical object as well as a psychological one. When we perceive we always perceive something and when we love we always love somebody or something. Consciousness can be understood only if we take into account the objective correlate, that is, the perceived, the beloved etc. The multifarious variety of real and possible objects of consciousness tells something about the specific nature of human consciousness. An object can be experienced as physical such as raindrops on my hand or as being of a non-physical nature like the hope for liberty and equality. They are both intentional objects and therefore part of the structures of subjective experience. And again, these structures reveal the complex nature of consciousness to be both physical and non-physical. Employing the Husserlian concept of intentionality in his analyses, Ricoeur immediately rejects two other theories of subjectivity: naturalism and subjective idealism (FN 9-10/13-14).

The naturalistic interpretation of subjectivity finds its expression in the methodology adopted by the empirical sciences. Only that which influences me causally is said to be part of the explanation of conscious experience, e.g. raindrops on my hand. The intra-mental experience, such as the hope for liberty and equality, is excluded from this kind of approach to subjectivity. The subject has to be treated in the same way as every other living being in nature. By contrast, the subjective idealism is an immanent conception of reality where the external, mind-independent world is considered if not irrelevant then at least as secondary for an analysis of subjectivity. There is no ontological difference between physical and non-physical entities, since both of them are part of my intentional consciousness. Thus, a hope for liberty and equality exists in the same manner as raindrops on my hands. A lack of differentiation between the physical and non-physical can, in this way, easily lead to the idea of an omnipotent subject who not only experiences the external world but creates it at the same time (FN 54/53, 485-6/456).

Ricoeur distances himself from such theories by indicating that '[t]he Cogito is internally shattered' (FN 14/17. Translation modified). The subject is made up by both physical and non-physical entities and therefore it has to be analyzed with a methodology different from the one used in empirical sciences without, however, sliding into an introspective mysticism: 'Inversely, knowledge of subjectivity cannot be reduced to introspection […] Its essence is to respect the originality of the Cogito as a cluster of the subject's intentional acts' (FN 10/14). The concept of intentionality

reveals the conscious subject as a peculiar being in the natural order. It belongs to nature and yet it transcends the immediate presence of sense-data. To approach a complex nature as that of human beings, we must find a method that includes as many aspects as possible of this particular being. Ricoeur finds such a methodology in a development of the two concepts inherited from Kant and Husserl. In the concept of intentionality he discovers an argument for the two-fold nature of subjectivity, quoting Maine de Brian's famous dictum in both works 'Homo simplex in vitalitate duplex in humanitate' (FN 228/213; FM 91/107). This two-fold nature excludes a method that is based only on causality, and by using the Husserlian concept in his analysis he is, at the same time, able to curb a subjective idealism. Once this is settled, he takes on the Kantian concept of transcendental synthesis in order to develop a basic notion of subjectivity.

Intentionality reveals the nature of subjectivity by turning the subject back upon it self. Ricoeur's favorite example to clarify this movement is the nature of perception (FM 26-7/44-45; NA 307-311/381-85; LS 256-7). When we perceive an object we are always bound to a certain perspective. The object is always seen from a certain perspective, from the front or the back; we are never capable of seeing the object in its totality. Perspectivity accounts for the limits and general narrowness of human perception. The conception of the perceived object points back to the nature of the perceiving subject as confined to a certain perspective. However, our understanding of the object is not completed in the act of perception. The awareness of our confined vision of the object is a result of the idea of other possible perspectives on the object. This knowledge somehow transcends our own perspective and reveals an understanding of the totality of perspectives on the object. This totality is the meaning of the perceived object. If I, for example, look at a plastic chair at the end of a long table, I only see a small part of it, say the top of the chair-back. Nevertheless, I immediately form an idea of that piece of matter as being part of a thing that we name a chair. This capacity to form an idea of totality from a fragmentary perception is what Ricoeur calls our 'power to express a meaning' or 'intention to signify' (FM 26/44). The subject is capable of talking about the object in its totality, and that so even in its absence. In this way, the understanding of the object as a thing reveals not only the restricted nature of human perception, but also to the human capacity to transcend that restriction in the act of expressing a meaning.

The subject is open to the object through its bodily senses, but in the act of understanding the object as a thing, it is turned back upon itself and revealed as more than just restricted perception. Thus, the subject is constituted by both receptivity (body) and spontaneity (reason), and any attempt to explain subjectivity must take into consideration both these aspects. Hence, the descriptive

analyses of intentionality in FN emphasize the necessity of a further analysis into this complex nature of subjectivity, which is then done in FM.

As said above, the two concepts have their specific development in the single works, intentionality is applied to the will in FN and the transcendental synthesis is used to analyze reason, sensibility and affectivity FM. To draw a systematic picture of the so-named stripped notion of subjectivity, I use the transcendental synthesis developed in FM as the basic structure by means of which the different constitutive structures of the subject shall be analyzed. The analyses of these structures draw on both works interchangeably. FM, however, is fundamental because of the effort to describe the structures of a basic human ontology in form of a philosophical anthropology. Here we find the mature Ricoeur arguing for a coherent notion of subjectivity. The work is, in many ways, the finest theoretical argument for his theory of subjectivity. It is, however, a very dense analysis, sometimes on the verge of obscurity, and in order to follow his arguments and transitions, we have to turn to the analyses in FN and, at times, the minor articles from the period. The intentional analyses uncover the peculiar structures and capacities of subjectivity (FN 8/12; PW 61); nevertheless, it is by means of the transcendental synthesis that these structures are put together to form unity we know as the human subject.

The Originating Affirmation

Now, we come to the last of the three fundamental methodological concepts, the originating affirmation. This concept plays a strange inarticulate role in the analyses. It is only dealt with explicitly in concluding pages of FM and in two minor articles, and still it seems appear whenever Ricoeur takes a stand on what is crucially human. It is employed more like a style than a working concept in his analyses. In all its actions, the human subject is characterized by an original will to exist: 'This will is no longer the object of statistical inquires, of inductive generalization; rather, it is recaptured as the primitive act of consciousness' (PW 66. Translation slightly modified; cf. TFA 291-2/362). The affirmation is primitive in the sense that it is a part of the nature the subject, and not a production of its activity. It is that which enables the subject to act in the first place, and thus, it has to be considered as an inherent feature of subjectivity.

To get an idea of how this concept is the major works, we have to look on two articles from the period in between FN and FM: *True and False Anguish* (1953) and *Negativity and Primary*

Affirmation[3] (1956). Ricoeur borrows the concept from another of his main sources of inspiration, the French philosopher Jean Nabert. Three quotes give a preliminary idea of what Ricoeur means by the concept: 'Under the pressure of the negative, of negative experiences, we must reachieve a notion of being which is *act* rather than *form*, living affirmation, the power of existing and of making exist'; 'philosophical reflexion is purifying in this: that it discerns the nucleus of affirmation shrouded in anger, the generosity concealed in the implicit will of murder' (NA 328/405; 323/400); 'As anguish becomes more radical, reflection will also become profound and will bring into view what I shall constantly be calling throughout the course of this meditation the "originating affirmation", something which we shall also attempt to capture at its successive levels' (TFA 288/358). The first quote indicates that Ricoeur's philosophy is existentialist in sense that it aims at understanding the problems of the living subject. It is not interested in the metaphysical forms or the essences of eternal truths, but concerned with how the human subject acts and exists in its concrete present. Ricoeur is, indeed, interested in metaphysical and ontological questions, but it is always with the living subject at the center of the analyses. His philosophy is a descriptive metaphysics of the acting subject with the intention to develop a specific ontology of subjectivity. The second and third disclose his conception of philosophical reflection. It must be optimistic and therapeutic in nature and, therefore, try to find the affirmative core underneath the perverse and malicious forms of human expression (ST 194). This emphasis on the affirmative character of the philosophical analysis is crucial, since reflects a pre-phenomenological conception of human nature. Human nature is not one with its apparently negative expressions, such as anguish, pain or malice. Of course, such expressions are to be considered as part of the subject, but, as it is said in the third quote, the more profound the expressions of negativity are, the more reflection to search even deeper for the originating affirmation. In this way, Ricoeur is glaringly clear about the nature, or, as he puts it himself, the *style* of his philosophy: 'a style of yes and not a style of "no," and perhaps even a style characterized by joy and not by anguish' (NA 305/378); and further in FN: 'Philosophy for us is a meditation of the *yes*, and not a surly intensification of the *no*' (FN 446/419).

The human subject is characterized by this primitive will-to-live (vouloir-vivre) or power to exist (puissance d'exister); however, Ricœur insists that this is not to be confused with a biological instinct, but, like the indefinite anguish in front of the possibility of death, the will-to-live is something more than a will to survive, and even more basic than rationality (TFA 290-1/361-362).

[3] The original title of this essay is *Negativité et affirmation originaire*. The term *affirmation originaire* is translated differently by the different translators (originating affirmation, original affirmation or primary affirmation). I will constantly use *originating affirmation*, since, in my opinion, this translation captures the original meaning best.

Being in the world, as mentioned above, is characterized by a disharmony between the two-fold nature of the subject (receptivity and spontaneity). The cogito is shattered, and subjectivity is in a state of conflict (FM 141/157). The original affirmation is a presupposition in his analyses, a conviction about human nature, and, as we shall se later on, it remains the fundament upon which he develops his theory of the human subject. Somehow, I believe, his entire theory of subjectivity is an attempt to argue for the existence of the affirmative character of humanity. That is why he repeatedly characterizes the nature of his philosophy as a recovery, restoration or reconciliation of the conflict within the subject (FN 17-20/20-23, 468-9/440-1; NA 306/379, 404; TFA 303/375).

This relationship between the originating affirmation and the nature of his philosophy is critical for understanding his analyses, because if we take the original affirmation to be an expression of something fundamental in the nature of the subject, other concepts such as verbe infini, conatus, appetitus, power to say and do, and attestation find a common point of reference. They all refer to an original will to say yes, an affirmation, a propensity to love and respect that go beyond the mere affectivity of the present. The originating affirmation, however, is in continuous conflict with the 'existential difference or existential negation' inherent in embodied nature of subjectivity (FM 135-8/152-4). The body tends to dim and pervert the fundamental will to be more than the present, more than the given reality (NA 327/404). Other than a propensity to be more than situated in a given present, to go beyond receptivity, the subject is fragile in the sense that it also tends to dispose of its affirmative character and hide itself in a sort of passivity. Instead of reflecting and acting on the immediate present, it tends to suffer the present. Giving into the overwhelming presence of the world, it is possible for the subject to blend into the world and become paralyzed in a state of passivity. As he writes about the nature of human habits, it is 'as if habit were a week point offered to what is perhaps the most perfidious of passions, the passion to become a thing' (FN 297/280). This passion to become a thing is an ever-present danger in the human synthesis. It is a way of disposing of spontaneity and becoming one with receptivity in order to relieve oneself of the responsibility that characterizes the human capacity to act.

However, this is not to suggest that Ricoeur operates with a dichotomy where, on the one hand, spontaneity, originating affirmation and reason are that which is essentially human, and on the other, receptivity, negativity and sensibility are characterized as antagonists to the humanity of the subject. The body, the experience of negativity and the concepts of habit and character are fundamental to subjectivity. In fact, they are all essentially aspects of the human subject, but their interplay in forming the subject is fragile and can very easily lead to a disproportion. This

disproportion is not only on the side of receptivity or sensibility, it is just as common to try to disregard one's body and rely solely on the powers of the mind. Such an attitude forces the subject into an equally dangerous state of disproportion, since it ignores the necessity to interact with the surrounding world. The subject is open to the world though the body, which remains an essential feature of its nature. The originating affirmation is embodied, as is the subject itself, and an account of subjectivity has to start from this fact of human nature. This is why philosophy has to be therapeutic in nature. It must set the subject right when it goes astray and falls into a disproportion, and in order to deal with negativity, immediate as well as reflected, it has to have an idea(l) of how the subject should be. Ricoeur's philosophy has such an idea(l), a pre-philosophical presupposition if one will. The primordial nature of subjectivity is affirmative, and this affirmative core is more basic than negativity in all its forms such as bodily limited perspective, character and vital feelings (FM 135-6/152). The positive style of his philosophy and the affirmative core of subjectivity are inseparably linked together throughout the entire development of his theory of subjectivity. They are the product of a conviction about human nature, and one could surmise that the aim of his philosophy might very well be to argue for the plausibility for this conviction.

The originating affirmation is the dynamic feature that drives the subject to engage itself with the world and the other human being. This interaction with alterity, the encounter with, reflection upon, and dialoguing with the world and the other human being, is made possible by this basic feature in subjectivity. In this way, we shall follow the development of the originating affirmation from its manifestation in the verb (vouloir-dire) over the notion of happiness to its most intimate expression in the feeling of happiness in the heart of the subject.

First, I will look on how consciousness is constituted by the transcendental synthesis of reason and sensibility. Then, this synthesis is put in relation to the world. The embodied subject is analyzed in its being embedded in the world and constituted by the interaction with this world and the other subjects in it. Finally, following Ricoeur's own emphasis on affectivity and emotion as crucial in subjectivity, I try to unravel the notion of emotions in the constitution of the subject and the role that they play in the subject's relation to itself, the world and the other subjects. This structure reflects Ricoeur's own structure in FM, but it reformulates the arguments by using the analyses of FN; thereby, in my opinion, the structural notion of subjectivity is clarified and enriched in a way not possible, or at least very difficult, by a reading of only one of the works.

Experience: Between Reason and Sensibility

The human subject is a complex being made up of reason and sensibility. This is a common intuition about human nature. Through our five senses we experience the world in which we live and with reason we can produce ideas and images without having to rely on immediate stimuli from the external world. How do these two basic capacities of the subject interact? Throughout the philosophical tradition there have been innumerous answers to questions of this sort, but by and large we can gather these answers into three generic groups: *dualism*, *materialism*, and *idealism*[4]. An extensive and central part of Ricoeur's writings deals with problems concerning the complexity of human nature, and one of his primary aims, from the earliest to the latest writings, is to give an account of the relation between body and reason in the constitution of subjectivity.

Ricoeur is very clear about his methodology. The methodology of scientific objectivity cannot explain the nature of the human subject: 'The body of a subject and the body as anonymous empirical object does not coincide. We can superimpose two objects, but not a dimension of the Cogito and an object. The lived body corresponds to a "behavior" of the will […] The dependence of my body on my self which wills in it and through it has nothing corresponding to it in the universe of empirical science except a body explained in terms of other' (FN 12/15. Translation slightly modified; cf. 207-8/193-4). Ricoeur is obviously not a materialist. The body, as lived and experienced subjectively, does not coincide with the physical body of the empirical sciences. However, what makes this statement particularly interesting is not the rejection of a pure causal explanation of subjectivity, but the focus on the body and not on the mind in this rejection. It is characteristic of Ricoeur's writings that even though his philosophy dismisses materialism, it never becomes neither idealistic nor dualistic.

What is in question here is not the nature of human reason or body, but the relation between the two aspects of subjectivity, embodied reason. The human body is not to be considered as an object

[4] *Dualism* is the philosophical position that holds that the human subject is made up by two ontological different parts of which one is spiritual and the other physical. These parts are heterogeneous entities that may communicate with one another but can never be explored by one and the same methodology. The philosophers that advocate *materialism* are certain that everything, when analyzed properly, turns out to be of physical nature. The experience of something non-physical is a phenomenological result of hidden, or at lest unknown, physical mechanisms. There are several degrees of materialism, giong from strong materialism (e.g. disappearance-theory/eliminative materialism that bans all psychological experience as illusionary or epiphenomena without any effect on the physical constitution) to weak materialism (e.g. translation-theory that defends the reality of the mental and the possibility of translating this into physical term, without disregarding the reality of neither of them). Finally, *idealism* claims that we can be sure only of the subjective experiences going on in our head. The existence of the external world is not a certain fact but may very well be a product of our mental operations. The world and the phenomenological experience of ourselves may only be a shadow of our real idealistic essence (Plato), or, as a contemporary philosopher puts it, we may be brains in a vat (Putnam).

among others in the world; it is a body lived by the subject in first-person. And it is somehow controlled or, at least, influenced by the will. However, regardless of his emphasis on the body, Ricoeur is clear about the status of the subject's two basic capacities. The body, being part of the involuntary aspect of subjectivity as opposed the voluntary, always has to be understood from the perspective of the conscious mind: 'seul un être de raison est aussi un être de sentiment' (LS 320). The subject is a complex being because it consists of both physical and non-psychical entities, body and reason. It is different from both angels (if one believes in such creature of pure reason) and the rest of nature's animals because of this complexity (FM 3/23).

The first, but not sufficient, step in an attempt to understand how reason and body interact in subjectivity has to deal with the concept of transcendental synthesis (FM 5/25, 17/35, 46/63, 49-50/67, 81-2/97-8). As we have seen, the transcendental synthesis is what enables subjective experience in the first place; it is that which brings together the heterogeneous capacities of the subject in what we know as consciousness. Although he quite faithfully adopts the concept of transcendental synthesis from Kant, Ricoeur himself believes that he makes a slight, but important, modification. Whereas Kant works out the concept in order to justify the objectivity of subjective experience and thereby emphasizes the scientific aspect of experience, Ricoeur is more interested in the subjective dimension of experience: 'namely that property of being thrown before me, at once given to my point of view *and* capable of being communicated, in a language comprehensible by any rational being' (FM 39/56)[5]. He describes this subjective feature of the object for 'the thing's objectival character'. By *objectival*, he refers to something that is more objective, or at least more fundamental, than scientific objectivity because it is that on which sciences builds, namely the pre-conceptual giveness of the world. The world as it is conferred to us through the senses and elaborated by reason: 'here consciousness is nothing else than that which stipulates that a thing is a thing only if it is in accordance with this synthetic constitution, if it can appear *and* be expressed, if it can affect me in my finitude *and* lend itself to the discourse of any rational being' (idem). It is important to notice that the experience made possible by the transcendental synthesis must be the

[5] The sentence deserves to be quoted in its full length, since it, albeit the dense and obscure formulations, presents a concise expression of a key point in his theory of subjectivity. He is concerned with the way the subject relates itself to the world around it, in short, the subjective experience. This exploration of subjectivity must begin, however, in the object, and not by means of some kind of introspective analysis. The sentence goes as follows: 'Kant reduced the scope of his discovery to the restricted dimensions of an epistemology. The objectivity of the object is reduced to the scientific aspect of objects belonging to a domain carved out by the history of the sciences. But criticism is more than epistemology, transcendental reflection is more than an exploration of the scientific nature of the objects of nature. The real *a priori* synthesis does not appear even in the first principles; it consists in the thing's objectival character (rather than objective, if objective means scientific), namely that property of being thrown before me, at once given to my point of view *and* capable of being communicated, in a language comprehensible by any rational being' (FM 38-9/56).

same for *any* rational being. This condition is what makes the synthesis transcendental in the sense that, even though it is what makes subjectivity possible, it is not a solipsistic kind of subjectivity, but the basic structures of subjective experience in every rational subject.

Therefore, besides distinguishing it from scientific analyses, the objectal character of the transcendental synthesis (effective experience of the world) is also a rejection of the introspective approach. Subjectivity must be examined by means of an analysis of its openness to the world (FM 19-20/37-8) and not by plunging into the hidden structures of individual subjectivity. The structures that Ricœur is looking for are the common structures involved in the way every reasonable subject is open to the world. Subjectivity is constituted in the encounter of reason and that which is not reason (that being the external world as well as the subject's own body). Neither reason nor body can account for subjectivity. Subjectivity has to be sought for in the interaction of the two. In order to understand how reason and body interact in the synthesis that forms consciousness, we have to turn to the concept of transcendental imagination (FM 5/25).

Transcendental Imagination[6]

TI is the cardinal concept in the synthesis because it is that which makes an integral union of body and reason. Ricoeur deals with two aspects of imagination, the transcendental and the empirical (FM 45/62). This distinction was first made by Kant in the *Critique of Pure Reason* where he assigns the transcendental imagination to the generation of the possibility (i.e. the basic structures) of experience, and the empirical imagination to the workings of actual experience (Kant 1997: 267-85/131-169). However, this distinction is not as rigid in Ricoeur as in Kant. As mentioned earlier, He is not interested in justifying the objective aspect of subjective experience, but concerned with the experience of lived subjectivity. He therefore continuously emphasizes that the transcendental is only the first step in the analysis. Then, before going to the next step, I shall take a closer look on the workings of TI.

TI is the 'blind point of knowledge' (FM 82/98) and 'remains obscure, hidden, and blind' (FM 45/62), and yet, we have to begin with this obscure concept in our analysis of the human subject (FM 5/25). It is in our encounter with the world that we first notice the function of TI. In the previous section, we saw that subjectivity is both receptive and spontaneous. Our bodily constitution permits only a limited perception of an object, a certain perspective. Our senses can never convey a complete perception, or what Ricoeur calls 'zero point' (FM 31/49). And still, we

[6] In the remaining of this section, I shall use the abbreviation TI for the transcendental imagination.

can transgress the limits of our point of view and, so to say, go beyond our perceptive abilities. We do this by judging about the fragmentary *presence* of the object in order to reach the *meaning* of the object. Ricoeur calls this human capacity for 'the intention to signify', and then adds 'to signify is to intend, the transgression of the point of view which is nothing else than speech as the possibility of expressing, and of expressing the point of view itself' (FM 26/44-45). He relates this 'intentional transgression' (idem) of our point of view with our will to say something about the perceived object, to give the object a significance. In this way, we discover two distinct capacities of the subject: the capacity to become affected by the world through our bodily senses, and the capacity, by means of reason, to confer a meaning to that which affects us.

To sum up what Ricoeur says about these two capacities of subjectivity, we can name them – as Ricoeur does himself – the *finite perspective* and the *infinite verb*. The finite perspective is what the bodily senses tell us about the objects in the world. Our bodily perception is always characterized by a fragmentary and sporadic registration of objects changing and moving in front of us. Thus, this registration is never complete, but linked to a certain perspective due to our bodily constitution. The infinite verb, however, is what permits us to form a meaning of this registration. We are able to say something about the object and categorize it as this or that, i.e., a specific thing loaded with a certain meaning. This capacity to categorize originates in our will-to-say-something about the object. It is an inherent feature of human subjectivity that distinguishes the human subject in the world: 'This transcendence of signification over perception, of speaking over perspective, is what makes the reflection on point of view as such possible: I am not immersed in the world to such an extent that I loose the aloofness of signifying, of intending, aloofness that is the principle of speech' (FM 31/48-49). The capacity to say-something-about the perceived object is rooted in the spontaneity of human reason, the infinite possibility of meaning contained in the verbal affirmation or negation of a certain object. In accordance with the previous example, we can affirm or deny that the object in front of us is actually a chair or not. This indicates something other than the bodily senses. However, neither of these heterogeneous aspects of subjectivity is in itself a human subject. They are heterogeneous in that senses are receptive and reason is spontaneous, but both are constitutional of subjectivity. Reason without the body and its senses is not a human reason. And a body without the mind is only a congregation of meaningless limbs and blind senses. We are made up of both mind and body, and in order to get them to interact in the single unity that we call the subject, we need a third concept, namely TI (FM 37/55).

TI is not given in itself and, therefore, it cannot be understood by scrutinizing in the subject, but must be inferred from the object encountered in the world through the senses and understood by reason (idem). The object tells us something about the way subjectivity is structured. The object is both registered sensibly and comprehended as possessing a meaning; it is not just an object but a certain thing. The object becomes a thing through the cooperation of body and reason; without this cooperation we only sense an undefined object without being able to determine and, thereby, understand that which moves in front of us or smells in the vicinity. That is why Ricoeur defines consciousness as an a priori transcendental synthesis and specifies that the nature of this synthesis is objectival (FM 38-9/56). It is a priori and transcendental because it is that which enables conscious experience of the objects in the world, and it is objectival since it is only understood through the perceived object, and not in itself (e.g. by introspection). But how does TI bring the two heterogeneous capacities of subjectivity together in the synthesis that we understand as consciousness? Ricoeur does not give a clear answer to this question, but joins Kant in stating that it remains an enigma (FM 41/59). Nevertheless, he relentlessly emphasizes that it is a fundamental concept in his theory of subjectivity because it remains the first step in an understanding of the subject (FM 5/25, 46/63). And although he does not give a clear answer, he attempts to close in on an approximate understanding of the concept.

It must be homogeneous with both sensibility and reason in order to mediate between those two capacities of the subject. Furthermore, it has to confer an identity to the subject through its different experiences. As Kant, he emphasizes that TI is the 'common root' of reason and sensibility (FM 43/61). What can satisfy these criteria? Ricoeur suggests that Kant himself pointed towards time as a possible clarification of TI (FM 42/59; cf. Kant 1997: 273-75/181-187) and then he follows that direction himself. Time is produced through my bodily senses, in my apprehension of the changing objects in front of me and the changing of my own body. Moreover, I use time to determine the different objects in relation to one another and to myself. Thereby, time is both sensible (forever changing) and rational (the relation to a non-changing identity). Time is not possible without a firm point that is not changing. This firm point is reason, e.g. our capacity to determine object as this or that, before or after, big or small, growing or decreasing; in short, our capacity to categorize. Our reason is the 'zero point' that we cannot perceive, but only reflect on (FM 31/49). Thereby, time mediates between sensibility and reason and procures us with consciousness of an identical subject and ever-changing objects. Time, however, does not explain the enigma of TI: 'Strictly speaking, we have only sharpened the paradox by a more subtle approximation […] To say that time is the

unity of that duality is to name the difficulty as well as to situate it – which is not nothing – but it does not solve it' (FM 43/60). This point of convergence between TI and time cannot be underestimated in Ricoeur's theory. Time itself remains an enigma like TI. For Kant, TI produces the transcendental schemes (Kant 1997: 273-4/180) which are in themselves nothing else than a priori time-determinations according to certain rules (Kant 1997: 275-6/180-1). TI thereby discloses the world as we know it. The transcendental schemes are the conditions on which the world appears for us. We do not perceive a chaos of ever-changing, undefined objects without a beginning, an end or an observable cause. On the contrary, the world appears to obey certain definable rules and follow the discrete, continuous succession of time. The world is both in continuous change and identical to itself. This, our understanding of the world, reveals our own, transcendental constitution (FM 20-1/39).

What is important for the present analysis is not to insist on the obscure conceptual architecture present in Kant and followed by Ricœur, in pages no less dense and inconclusive, but to concentrate on the convergence itself. Why does time sharpen the paradox inherent in TI?

If TI is the hidden, but 'common root' in subjectivity, i.e., that which procures the identity of the subject, and if time converges on this root, then this tells us something fundamental about the nature of subjectivity: subjectivity is in its core marked by temporality, this is, '[m]y life is temporality' (FN 425/399-400). Our understanding of the world and, thereby, of ourselves is temporal (an insight that remains and is further developed in the later writings).

In order to understand this basic temporality, we have to turn to the analyses of imagination in FN. As I said earlier, the distinction between transcendental and empiric is not as clear-cut in Ricoeur as it is in Kant, since Ricoeur is concerned with the structures of experience in the living subject. Then, we may clarify the relationship between TI and time by looking at the workings of the empirical (i.e. concrete) use of imagination.

A central passage sums up the analyses of imagination in FN: 'But the core of meaning of imagination remains knowledge […] It is as knowledge that imagination which swells our desires is susceptible to coming under the control of the will and that our life itself can be evaluated' (FN 99/95. Translation slightly modified). Imagination procures knowledge about the world and the subject by creating an evaluative space of reason. Ricoeur compares it to light: 'Light is thus the space of appearance, but light is also a space of intelligibility. Light, as openness, is a medium of appearance and expressibility' (FM 40/58). We are immediately aware of the workings of imagination in our understanding of the world and our self. Imagination is reproductive: a

representation of an absent thing or situation, re-presenting the thing or the situation that was previous the object of our conscious awareness (memory of the past). On the other hand, it is productive: an image of a possible, or impossible, thing or situation that has not yet taken place. We can imagine a scenario and work towards it (a possible future). Put like this, we now begin to see how time and imagination can help to explain one another. One can say, in a very simplified way, that our capacity to categorize an object is founded on our capacity to re-present past perceptions, compare them to one another, and, then, make a judgment about what is going to happen in the future. Thus, it is through our imaginative powers that we come to understand about the world and ourselves, in short: imagination is the basis for knowledge (FN 99/95). And the way imagination procures knowledge shows that this knowledge is always temporal. We come to an understanding of the workings of the world through our capacity to re-present the past and project the future in our imagination.

I noticed that Ricoeur maintained, although in a less strict sense, the obscure Kantian division of the transcendental and empirical aspect of imagination because it tells us something important about his notion of subjectivity. Imagination is not just something we use in our dealings with a pre-given, objective reality. As if first we have the reality as it is in itself (objective, firm, inescapable) and then, in a second step, we have our own interpretation of that reality. This can be a first-hand, intuitive understanding of the relation between world and the perceiving subject. But often the 'everyday forms' of knowledge tend to hide or complicate 'man's structures or *fundamental possibilities*' (FN 3/7). Therefore, human reality might more complex than assumed by such a naïve realism.

The empirical aspect of imagination is subjective in the sense that every subject uses it in its own way. How we employ it depends on our personal and social history, individual needs, private dreams and ideas. This we all know. The transcendental aspect, however, works in the same way in every subject. And this is far less intuitive. It is that which brings together reason and sensibility in what we know as consciousness. The most basic structure of subjectivity is a synthesis of our receptive capacities, the bodily senses, which let us become affected by the world, and our spontaneous capacities, our infinite verb, our will-to-say, our reason which enable us to affirm or deny and, thereby, confer a meaning to the object. This synthesis (i.e. consciousness) is made possible by TI in that it mediates between the two heterogeneous aspects of subjectivity. As with the workings of the empirical imagination, TI creates the evaluative space of reason that we call reality. Human reality is a product of the external stimuli received through the senses and the

affirmative powers of reason. Now, this may sound like subjective idealism in which reality is a product of the subjective mind. This, however, is not the case. Human reality is not a product of mind alone. As stated repeatedly throughout the analysis, reality made up of both body and mind. Intuitively, we see the external word as something separated from the perceiving subject. And the world is indeed external and accessible only through the senses. Our body is the organ for our actions and recipient for (some of) our sufferings. Thus, subjectivity is only understood as embodied reason. And as Ricoeur puts it: 'To succeed here we have to reintroduce the body into the cogito as a whole and to recover the fundamental certitude of being incarnate, of being in a corporal situation' (FN 217/203; 87-8/84-5). What Ricoeur attempts to show is that even in its most basic experiences, such as experiencing time and the external world, the subject is, somehow, active in that it confers a meaning to that which affects it. The subject is active in experiencing the world and itself because it is judging about what affects it. Human reality is loaded with meaning, and the subject is, somehow, a co-producer of this meaningful reality. This is why TI is so important in Ricoeur's theory of subjectivity. It shows that a meaningful reality cannot rise from reason alone. Our conscious experience is made possible by the mediation produced by TI, since it procures the unity of our experience of the world. TI shows how reality is both mental and physical in that reason's ability to determine and comprehend is always conditioned by the external stimuli. We are affected by a wrong determination of a certain object. This can be illustrated by an over-simplified example. Say we affirm that the object moving toward us is a leaf and not a baseball, and we go about, undisturbed, with whatever we were doing. Less than a second later, we are definitely affected by our wrong affirmation. This is not true of only leaves and baseballs, but of any object of our experience, from the hard physical objects to the most feeble passion moving in our breast.

Surely, Ricoeur would not deny the presence of an independent physical universe in which the subject is but a small actor in a tiny, fragile atmosphere. This, however, is not the subject-matter of his analyses. What he is trying to demonstrate is how we can access the peculiar nature of the human subject and understand how this subject perceives, feels, acts and suffers. This section has been an analysis of how reason and sensibility are interconnected in the transcendental (i.e. valid for every rational human subject) synthesis that we apprehend as conscious experience. This is a necessary, but not sufficient, first step in his theory of subjectivity: 'Consciousness is not yet the unity of a person in itself and for itself; it is not one person; it is not a person' (FM 45-6/63. Translation slightly modified). I have argued that even the most basic experience, as that of reality, is a complex matter. The subject is experiencing the world around it by means of a capacity to judge

about that which affects it (external stimuli as well as time and the subject itself). Ricoeur names this capacity 'the infinite verb', 'will-to-say', 'power-to-affirm'. Affirmation, however, is closely linked to the possibility of failure (FM 32/50). We can make wrong judgments about the world and ourselves. Thus, Ricoeur introduces two notions that have large impact on his further analysis, namely the concept of disproportion and the concept of human frailty. I return to these notions in the last section. In the following section, I shall look on how the subject is able to act in the world disclosed by reason and sensibility.

Action: The Practical Space of Reason

In the previous section, we saw how sensibility and reason are related in the making of consciousness. This, however, is only an abstract conception of consciousness; these are the basic structures of subjective experience, but not sufficient for self-consciousness and, therefore, not yet a person (FM 45-6/63). How, then, does the subject become conscious of itself and thereby a person? In order to answer that question, we have to go beyond the transcendental synthesis and look at what I have named the practical space of reason. The transcendental analysis is necessary because it discloses the most basic structures of conscious experience and, further, explains the integrate workings of the two fundamental aspects of this peculiar experience, namely body and reason. But still, this approach has its limit 'because the universe of things is still only the abstract framework of our life-world' (FM 47/65; 106-7/123). The living subject does not only experience the world, but is interacting with that reality. The transcendental analysis clarified the basic structures that enable the objectal experience that we know as the world. The analysis of the practical space of reason goes a step further and attempts to disclose the structures by means of which the subject copes with that world. If the former analyses enquired into how the subject perceives and understands that which it experiences as reality, this section deals with how the subject interacts with this reality in terms of decision, motivation, and value, and how these come together in what we understand as human action in the world.

The transition from the transcendental to the practical is marked by the concept of volition. The reason why the subject analyzed by means of the concept of transcendental synthesis was not yet conscious of itself is because those analyses focused only on the structures of the experiencing subject. However, it is only by interacting with the experienced world that the subject becomes conscious of itself as an individual subject, an agent; or as Ricoeur puts it: 'the self is in its acts' (FN 58/56; cf. FM 51/68), and further, 'I affirm myself as the subject precisely in the object of my

willing' (FN 60/58). By acting, the subject thickens the concept of identity found in the transcendental analysis. The concept of identity through changes pertains to both animate as well as unanimated beings, whereas the concept of subjectivity requires something more than mere numeric identity to differentiate itself from other entities in the world. Ricoeur finds this in the concept of volition: a peculiar human capacity to cause a change in the external world without being directly induced to this action by a prior external cause. If the previous section showed that the scientific method of cause and effect could not explain subjective experience of reality, this becomes even clearer in the present analysis. The function of human will is different from the physical explanation of cause and effect because '[c]onsciousness is not a natural phenomenon' (FN 67; cf. 12/15-6, 187/177-8, 379/372-3). The will is one of the critical capacities that differentiate human subjects from other beings in the world. Therefore, to understand human action, we have to clarify the concept of volition, and this is possible only through an analysis of the interaction of decision, motivation and value.

In an important chapter in FN, Ricoeur draws a poignant picture of both human will and the peculiar ontology pertaining to subjectivity: 'A thing is here, located, and determined by what is not itself, while freedom is not located, does not become aware of itself, does not discover itself as being already here until I observe it. It creates itself in doing and affirms itself to the extent to which it does – it is a being which determines itself. Its potential being is not at all a gasping abyss, it is the actual task which freedom is for itself in the moment *in which it constitutes itself* by the decision it makes' (FN 65/63). Whereas other physical objects in the world are determined by their presence, their being-there as an effect of a prior physical cause (and are themselves a cause for a posterior effect), the same cannot be said of the subject. The subject has a will that is not an effect of a prior physical cause, or at least not entirely an effect of a physical cause. There is something more than external stimuli to the being-there of subject. It is not just a link in the great chain of causality. When the subject reacts to external stimuli, then the effects escape the scheme drawn by causality. In short, the output cannot be explained by a causal deduction of the input (FN 94-5/91). The input passes through the practical space of reason of the individual subject before it gives rise to certain effect. It is at this point that the will appears.

Conscious experience is a result of our affirmation and categorization of the perceived objects as specific things. It is in this actual experience that we come to notice the workings of the will. The will is not present as something we can understand and then use. On the contrary, we come to know the will through its works, i.e., the decision to act. Through a reflection on our decision, we can

come to know what made *me* (*I myself*, this individual subject in a certain situation at a given time in the context of the world) chose that particular decision and not another. But reflection comes afterwards. Therefore, it is 'the awareness of that practical reference to myself which is the very root of reflection. Explicit reflection, in the distended form of the judgment "it is I who…," only raises a more primitive affirmation of myself to the dignity of discourse, a judgment which projects itself in the plan of action' (FN 61/59). Once again, we encounter the concept of affirmation. There exists an affirmation more primitive than the subject itself, and every conscious choice that the subject makes originates in this primitive affirmation inherent in subjectivity. I shall come back to this in the last section.

In the present context, we must concentrate on the relation between reflection and the practical self. Before its capacity to reflect, the subject is characterized by its capacity to act. And although reflection and action are inseparable in every aspect of the actual voluntary act, the distinction is interesting because it emphasizes that the subject is only a subject through its actions. The subject is not only a *res cogitans*, but first of all a *cogito agens*. Reflection may be the hallmark of humanity. Nevertheless, reflection is always rooted in sensibility and action. The subject senses, thinks, suffers, and acts. Thus, Ricoeur coins the term *Integral Cogito* to describe the nature of the general human subject (FN 91/88). The term 'integral' is curiously enough omitted by the translator). The subject is not given before its actions but exists only in virtue of its actions: 'I am in a world in which there is something to be done. I have embarked into it in order to act in it' (FN 212/197-8); or simply, '[t]o exist is to act' (FN 334/316).

What is important, in this section, is not so much the nature of action itself (which will be dealt with in chapter two), but *how* the subject is capable to act and *why* it acts in a certain way. How it acts is analyzed by the concept of decision; and why it acts in one this way and not in another is approached through the concepts of motivation and value. The three concepts are obviously inseparable in effective action (FN 71-2/69, 83/80, 93-4/90), and they shall, therefore, be analyzed in their reciprocal workings. Nevertheless, for the sake of clarity it might be a good idea to start with a brief presentation of Ricoeur's basic understanding of the three concepts.

Decision

Ricoeur gives a concise definition of decision: '*a decision signifies, that is, designates in emptiness, a future action which depends on me and which is within my power*' (FN 43/42. Translation slightly modified). Decision is constituted by intentionality. It is a decision to do something (a future

action). And intentionality is what makes decision a fundamental link between the world and the subject (FN 42/42). By acting, the subject interacts with the world; it places itself in the world, so to say. However, decision does not necessarily depend on the world. One can decide to do something not regarding or, perhaps even more significant, in spite of the world. The power to decide is essentially a capacity to designate in emptiness (désigner à vide), to conceive a project about what to do without taking anything into account except *oneself*. And this is, perhaps, the most critical feature of decision. To decide is, in the strictest sense, only dependent on the subject. And the individual subject is ultimately a result of its decisions: 'It depends on me to decide' (FN 184/175). This is a necessary first step is understanding of the concept of responsibility (FN 57/56, 81/78).

So, the subject affirms itself in the object of its will, and, in deciding on what to do with the object, it characterizes itself as an agent of certain actions. These actions are done by the subject and, therefore their meaning is partly created by the subject. And vice versa, the subject becomes this particular subject who has done such actions and thus becomes a meaningful subject by these actions. Ricoeur calls this feature of decision '[t]he imputation of myself: "*se* decider" – making up *my* mind' (FN 55/54).

Motivation and Value

However, no decision has ever been made without a certain motivation: 'There are no decision without motives' (FN 66/64). The concept of decision as the capacity to designate in emptiness, i.e., dependent only upon the subject itself, could easily lead to the erroneous conclusion that the human will is arbitrary. On the contrary, '[t]he highest form of will is the will which has *its* reasons, that is, one which bears at the same time the mark of my initiative and the mark of a legitimacy' (idem. Translation modified). Thus, when the subject makes a decision, then the decision determines the subject as this individual subject by referring to some kind of legitimacy for the decision made, and, consequently, for the action of the subject. It is by giving an account of the motives for a certain decision that the question of legitimacy comes into the picture. A motif can be legitimate or illegitimate. So, the question is: how does the subject understand what is a legitimate motif and what is not? To answer this question, we must take a closer look on the concept of value.

As pointed to several times now, the human subject is made up of both reason and body, and the two aspects are intrinsically interlaced in subjective experience. Nevertheless, it can be a help to examine the concept of motivation from these two aspects separately, i.e., the bodily motives and the motives of reason. This is, of course, a forced distinction, since motives never present

themselves as pure bodily motives or motive of pure reason. Anyhow, the theoretic distinction might clarify the structure of motivation.

The concept of motivation covers a wide and complex field of phenomena. But at the extreme poles of the concept, we find the bodily rooted motives at one end, and the motives of pure reason at the other. This scale goes from the absolute involuntary (body) to the absolute voluntary (reason). Thus, our decision is influenced by all kinds of motives ranging from those that we are necessarily affected by to those that we deliberately take upon us. I begin with the bodily motives.

It is the body that evidences the close connection between motivation and value: 'Before I will it, a value already appeals to me solely because I exist in flesh […] The mystery of incarnate Cogito ties willing to this first stratum of values with which motivation begins' (FN 94/90). Motivation is generated by values. Although we normally ascribe generation of values to the power of reason, Ricoeur makes it clear that values are strictly bound to the embodied mind, i.e., our existential situation of being both body and reason. So, there exist no values without bodily existence: 'The first non-deducible is the body as existing, life as value' (idem). It is through the body that we first feel a motivation to do something, to act. Ricoeur name this first layer of value 'organic values' (FN 120/114). The organic values are characterized as needs. We feel a motivation to do something in order to satisfy a certain need; thus, it is clear that we act not by reason alone, but are always affected by our existence being bodily existence (FN 86/83). Examples of basic human needs are thirst and hunger (idem). We are not in control of our needs. They present themselves as an involuntary part of our existence and demand to be satisfied. If we do not satisfy these basic needs pertaining to our existence as an integral Cogito, we simply languish and die. Hence, they are immediate motives for certain actions, e.g., eating or drinking. The question is how they become values? To answer this, we must turn to the motives of (pure) reason.

The subject cannot control its needs, but '[t]hough I am not the master of need in the sense of lack, I can reject it as a reason for action. In this extreme experience man shows his humanity' (FN 93/90). The subject can choose to not satisfy its needs. And, perhaps even more significantly, this choice can be done systematically. Some of the most dramatic examples of such choices are celibacy and hunger strike: 'Man is capable of choosing between his hunger *and* something else' (idem). It is a motivation to act in spite of the spontaneous motivation generated by bodily needs. In this way, we come about the strange situation in which the subject turns against its own body, its own existence, in order to follow another motivational factor. This human capacity to sacrifice its needs and, in the most extreme case, its life inflates the complexity of motives and values. The

human will can be controlled by something else than the body and '[n]eeds can thus be one motive *among* others' (FN 94/90). This is what I call a motivation of pure reason.

The connection of motives and values takes place in the motivational field disclosed by the two extreme poles, the bodily needs and the sacrifice of these needs by (pure) reason. A sacrifice by means of pure reason is very different from suicide: 'In contrast with suicide, which is pure negation and destruction, sacrifice is entirely affirmative, affirming value and being, albeit without reference to my life' (FN 121/116). Whereas suicide can be an expression of an immanent need in the individual subject, a sacrifice is a (extreme) way to transcend the proper needs (and proper motives) and partake in the practical space of values. The motivational staircase ascends from the involuntary bodily needs to the highest expression of voluntary decision, the sacrifice of one's own life. Of course, human decision is very rarely either a blind following of bodily needs or a dramatic sacrifice of our own lives. Most of our decisions are in a more moderate position between these poles. In the choice of following one motive instead of another, the concept of value is introduced into the analysis of motive and decision.

Bodily needs are not commonly understood as values in themselves, but more a necessary part of human existence. However, they become values by being only certain motives amongst others[7]. The human subject holds the possibility to decide amongst an infinite variety of different motives. Because of this possibility, the subject is seizing the different motives in terms of their value. It decides on one choice and not another due to the values that it ascribes to different motives: 'on the human level, organic life is undoubtedly a cluster of *heterogeneous* demands, revealing *discordant* values' (FN 105/101). The possibility to choose is constituted by the practical space of reason where the discordant and heterogeneous values are weighted against one another in order to decide on which to choose as a legitimate motive for a certain action. Thereby, the motives (even the spontaneous bodily ones) become a hierarchy of values for the individual subject (FN 145-6/138-9).

The question is now: how do the different motives form the hierarchy of heterogeneous values that structure the practical space of reason?

[7] Even though the bodily or organic needs are only some motives amongst others, Ricoeur is well aware of their fundamental character in that a need is 'the primordial spontaneity of the body; as such it originally and initially reveals values which set it apart from all other sources of motives. Through need, values *emerge* without my having posited them in my act-generating role: bread is good, wine is good' (FN 94/90).

The Practical Space of Reason

Before answering that question, we have to take a closer look at what I mean by the practical space of reason. In the previous section, we saw how the transcendental imagination made conscious experience possible by the synthesis of body and reason. This, however, cannot account for the subjective experience as being a part of and, at the same time, different from the experienced world, i.e., the experience of being an autonomous person somewhat different from the rest of the world. The subject does not blend into the world, but is itself capable of causing a change in the world through actions. And actions, on their part, are done because of a certain motivation. Motives do not compel the subject to act in a certain way, but present themselves as different options among which the subject can decide on which one to choose. These heterogeneous motives occur as options for the subject in what I have chosen to name the practical space of reason.

Whereas conscious experience is roughly the same in every normal functioning subject, the practical space of reason is personal, in the strictest sense of the word. Conscious experience is intersubjective because it is that which enables a common experience of objects in and features of the world. The practical space is personal because it is that through which subject decides on how to interact with the world. By means of these decisions, the subject determines and affirms itself as an agent of certain actions and, thereby, it becomes this particular subject, a person.

In this context, we once again encounter the concept of imagination. This time, though, it is not the transcendental, but the empirical aspect that is at play. Imagination is 'imagination of the absent thing and of action directed towards the thing – the crossroad of need and willing' (FN 95/91. Translation modified). Whereas transcendental imagination illuminates the basic structures of the experienced world, the empirical aspect is that which presents the motives (and the means to follow those motives) for the subject to decide on. The quoted sentence says it clearly: imagination is the crossroads of the will and the needs. However, imagination does not only present the means to fulfill bodily needs, such as thirst or hunger, but an infinite variety of motives. As we have seen, these motives span from the most concrete bodily need, rooted in our bodily constitution, to the most abstract motive, such as the decision not to eat in order to serve a motive different from the motive to fulfill the need of eating (i.e. an idea). The subject can decide to follow one motive and not another.

Imagination plays a critical role in decision-making. The capacity to re-present absent things, feelings, and past scenarios and to project imagined decisions into a possible future makes it the catalyst in the practical space of reason. Without it no decision would ever be made. It holds

together past, present and future in the process of decision. This is an important feature of imagination because it makes possible what we experience as identity through time. In re-presenting the past and projecting the future, it confers the experience of identity through a changing world. It is the same subject who has done such-and-such actions in the past and the same who will do these or other actions in the future: 'I impute a future action to myself in identifying the projected "I" with the "I" which projects' (FN 140/133).

Another feature of imagination, closely linked to the aspect of time, is its capacity to present a motive as connected with either pleasure or pain.[8] Every situation and every action is experienced by the subject as joined with a certain feeling. These feelings cover an infinite field of different sensations and emotions: pain, fear, tiredness, impatience, anguish, satisfaction, energy, joy, happiness etc. They are, in the first place, revealed by our bodily needs. If we do not eat, we become tired, impatient and end up being in pain. Therefore, Ricoeur insists that bodily needs are a 'primordial spontaneity' that generates values without the subject's consent. It is a pleasure to satisfy ones need, and we therefore spontaneously ascribe a positive value to the satisfaction of our needs: 'prereflexive feeling is a spontaneous belief concerning the good of the body' (FN 104/100). Imagination is what enables the subject to break with this spontaneous belief on the bodily good. Instead of immediately satisfying the bodily needs, these needs are configured in the practical space as a motive for an action and not an immediate, necessary cause. Of course, these needs still remain a strong motivation for action, since they promise a pleasurable satisfaction and thereby generate a desire for doing a specific action, such as eating. However, the subject can decide not to satisfy this need. This is possible because imagination, in configuring the practical space of reason, can present this motive together with other motives which perhaps are capable of generating other kinds of desires, such as not eating because of a desire to follow certain principle, or simply because of a desire to loose weight.

This is an interesting feature of human subjectivity. The feeling of pleasure or pain points to a rudimentary system of value inherent in subjectivity. A decision to follow one motive and not another in the practical space of reason is immediately joined with a certain feeling: 'It is the affective medium of all values: a value can reach me only as dignifying a motive, and no motive can incline me if it does not impress my sensibility' (FN 122/117). The heterogeneous and discordant motives present themselves in the practical space of reason where the subject experience

[8] This is a very clear-cut distinction that does not correspond to the effective experience of pleasure and pain: 'Affectivity forms no system; it illustrates disparate values in disparate pleasures and suffering' (FN 120/115). Anyway, I shall use this abstract distinction in order to clarify the structural connection between motives and values.

them and must decide which one to follow. Still, the subject would never come to a decision if some motives were not highlighted and others not. This is the job of feelings. They order the different motives into a hierarchy of values (FN 146/139). I return to the question of feelings in the last section.

This hierarchy, however, is not organized according to the immediate feelings of pleasure and pain. The immediate action upon a motive is an instinct. Most animals act according to instincts. This could be an instinct to eat, run, hide, attack, or reproduce. And some animals have almost perfect instincts, that is, they react immediately upon external stimuli. The human subject, on the contrary, has more instincts (in virtue of its elaborated time-consciousness and unlimited capacity to invent through knowledge), but is less instinctive in its actions (FN 94-5/91). Human decision is permeated by a degree of hesitation not figuring in other animals. Hesitation is an indication of freedom: 'Hesitation is a choice seeking itself' (FN 137/131. Translation modified). It discloses the fact that motives are not causes to actions, but always *possible* motives. The subject hesitates because it is uncertain about what to do in a particular situation: 'In indecision I am lost among confused motives' (FN 142/135). Motives become values through feelings. The human subject, though, is not only affected by the values pertaining to bodily needs. There is something other than bodily needs at play in the human hierarchy of values. Other than the fulfilling of mere desire, human action is characterized by a quest for legitimacy. The question of the legitimacy of a value is posed by the integral Cogito because it is both body and reason. The subject is capable of opposing its immediate desire to satisfy a bodily need.

Now we have the necessary features to understand the basic structures of the practical space of reason. In the experience of the external world, we affirm ourselves as a part of and, at the same time, as different from this world, by means of actions that are generated by certain decisions. These decisions are not directly compelled by stimuli from the world, but are, in the strictest sense, only dependent on the subject (to designate in emptiness). The individual subject becomes this particular person through its decisions and actions. No decision, however, is made without a motivation. There exists an infinite variety of heterogeneous and discordant motives among which the subject can decide, from bodily needs to abstract ideas. Motives are intrinsically joined with a certain feeling, since all external stimuli and every action are characterized by different feelings, roughly gathered within a scale of pleasure and pain. These feelings make up a rudimentary system of values: good = pleasure vs. bad = pain. Nevertheless, human decision is more complex than that. This complexity is due to what I call the practical space of reason.

Imagination configures an experiential space wherein the different motives and their joined feelings are weighted against one another according to a hierarchy of values. The configuration of that space is due to both past experience, the present context and the individual constitution of the subject. The individual constitution of the subject is determined by both body and reason. Not only satisfaction of bodily needs, but also ideas of reason present themselves as values in the practical space. These ideas are generated by past experiences, culture, upbringing, personal desires, bodily constitution etc. The central point is that they are formed by the individual subject and are experienced as values among other values in the practical space of reason. This can lead to a confusion of values that affects the subject's decisions and actions and thus the basic identity of the subject. The analysis of the concepts pertaining to volition, i.e., decision, motivation and value, has shown how the subject becomes an individual subject, a person. Personal identity is determined by the subject's actions. However, we have also seen that deciding on what to do, how to interact with the world and, in this way, to become a person is a highly complex matter. Personal identity is not a fact, but a problem. Values can solicit confusion and conflict in personal identity. By acting in a certain way, the subject consolidates itself as a person with such-and-such values: 'I consecrate the hierarchy of values precisely in choosing' (FN 174/165). And yet, an action can be in conflict with the subject's idea of itself as a self and solicit a conflict between the subject as self and the subject as person. I return to this problem in the last pages of this chapter.

The first section dealt with how the transcendental synthesis constitutes conscious experience. This experience is intersubjective, i.e., every normal functioning subject experience the world in roughly the same way. In this section, we have seen that this is not the case with the subject's interaction with the world. The subject's decision is strictly personal; in fact, it becomes this particular person by means of its decisions. Thus, the practical space of reason is personal (contrary to conscious experience). This raises the question of the legitimacy of the values that motivate the individual subject to make its decisions and act in the practical space of reason. How do autonomous and heterogeneous persons interact with one another in a common world? This question is the topic the next section where I shall look closer on the question of identity and legitimacy in front of the other person.

The Subject as a Person in Front of Other Persons

Human beings perceive the world in fairly the same way. Their world, however, is not the same. Although we possess the same senses and act by means of similar rational structures, every subject colors the experience of the world in its own personal way (FM 61/78). We interact by means of certain motives that reflect our values. These values are generated by our own nature, and by that which is different from us (the world and the other subjects). This reciprocal relation between subject and world takes place in what I have named the practical space of reason. In this space we form, through our decisions and actions, what we know as the individual subject, a person. In this way, the subject becomes a person through the way it interacts with the world. But how, if every subject copes with the world in its own particular way, do different persons interact with one another in a world to which they all belong? To answer this, we have to look at the values by which the subject acts and, in particular, on the legitimacy of these values.

We saw earlier that the highest will is 'one which bears at the same time the mark of my initiative and that of a legitimacy' (FN 66/64. Translation modified). These are two critical aspects of the concept of a person. The subject's initiative (spontaneity) is what makes autonomy possible and sets it free of the immediate response to stimuli. Since the subject is free and its actions, in the strictest sense, only depend upon itself, it is also responsible for these actions. But to whom or what is the subject responsible and how does it decide on which values to follow and which to turn down? These questions point to the second feature, namely the legitimacy of the subject's values. Earlier we saw how conscious experience develops into the particular experience of the individual subject in that the subject becomes a person through certain decisions and actions in the practical space of reason. But Ricoeur writes unambiguously that '[o]ne would be greatly mistaken, however in holding this synthesis for an accomplished one, given in itself in the immediacy of self to self. The person is still a projected synthesis that seizes itself in the representation of a task, of an ideal of what the person should be' (FM 69/86; cf. LS 261). The subject is not a person simply by the fact that it is an autonomous and free being. To be a person is not a fact, but an open task for the individual subject. But before going further into to the concept of the person, we first have to clarify the relation between legitimacy and responsibility.

Legitimacy becomes a critical issue when the subject encounters another subject with a hierarchy of values different from its own. If the subject were only to care about its own existence in its decisions and actions, legitimacy would be a much less complicated matter than it actually is. The subject would, in such a state, only have to order its hierarchy of values in order to promote its own

well-being[9]. But the subject is continuously exposed to the decisions and actions of other subjects. The meaning of the movements and gestures of the other subject is not understood by laws of causality, but indicate the workings of intentionality, i.e., a meaning-giving reason analogous to my own. Thus, we discover that: 'This subject is myself, yourself, it is my fellow man. My own experience of myself and the sympathy (or better, intropathy) for the other are two living experiences which give rise to phenomenological concepts immediately valid for subjectivity in general' (FN 226/212)[10]. Thus, we are induced into the complex dialectic of autopathy (the feeling of myself) and sympathy (the feeling of the other). What is important in the present context, however, is not so much the complexity of the presence of the other, but what this presence does to the values of the individual subject.

Why does a subject have to care about the values of another? Are we not only responsible of our own well-being (i.e. organize our hierarchy of values in order to promote such a state)? Such solipsistic perspective on the world would soon bring very the subject in difficulty. If only *my* values and *my* well-being count, then my identity shatters within a few moments. I live, de facto, in a human world constituted by other human subjects. My identity depends on other human subjects because it is an identity in front of 'my fellow man'[11]. We are similar and yet not the same. I am myself and not the other, and vice versa. I cannot pretend to know how the other understands the world. I can try to understand how she understands the world by ways of observation and imagination. And still, the existence of the other surpasses my imagination. A simple meeting with the other proves this point: 'the hand of the Other that I see "appresents to me' the solipsistic touching of that hand and all that goes along with this touching. A whole world is born to that hand, a world that I can only presentiate, "render" present to myself (vergegenwärtigen), without its being present to me' (APH 66/122).

The presence of the other subject's hand is critical for my identity because of the implications of this presence. It is a presence that I cannot be indifferent of. If my values do not regard the values of the subject revealed by the hand on my shoulder, I would not be able to interact in a human world. I would end up undermining my own identity. As a person my being is a social being, or a political animal. (FM 120/136). Thus, I must be interested in how the others consider me in order to exist in

[9] 'Pleasure, pain, the useful, the agreeable, and the easy, in spite of their heterogeneity, still have the feel of a family bound together by the term *well-being*' (FN 116/111)

[10] 'il appartient à l'apparence de la personne de se donner non seulement comme la présence d'un être, donc comme une apparence ontique, mais, en outre, comme l'apparence d'un autre sujet, l'apparence de mon semblable.' (SR 267; cf. SR 269; KH 196-8/247-48).

[11] 'by formulating concepts of the subjectivity derived from perception of the self and from understanding of the behavior of the other as second person' (FN 227/212. Translation slightly modified)

a world made up of a manifold of different subjects. To interact with the other, I have to recognize her as a person with her own values and desires (i.e. her own identity). As I stand before her, she stands before me. We are both persons, but with our own characters: 'Human likeness is implied in the mutual otherness of individual characters' (FM 70/88). Her presence demands me to consider her and recognize that she may have values and desires different from mine. In interaction with one another, there must be a reciprocal recognition of our respective identities. Therefore, my desire to promote my own well-being has to be limited by the simple fact that other subjects exist: 'en effet, je ne puis limiter mon désir en m'obligeant, sans poser le droit d'autrui à exister de quelque manière; réciproquement reconnaitre autrui, c'est m'obliger de quelque manière; obligation et existence d'autrui sont deux positions corrélatives.' (SR 274).

Hence, the identity of the subject is not rooted only in decisions and actions concerning its own desires, aversions, fears, joys, in short, its own wellbeing. The concept of identity hinges on the recognition of other subjects. Identity is constituted by actions, and, since actions inscribe themselves in a world made up by *this* subject *and* other subjects, they have to be done according to the others and myself. Put in another way, the subject is not only responsible to itself, but also to the other subjects around it. In fact, the concept of responsibility is coined through this being with the others. A subject is responsible because it must answer for its actions and decisions to itself *and* to the other subject whose mere presence obligates an answer on my part.

If the subject's identity is coined through its actions and the actions, on their behalf, are done according to certain values, then we can say that the subject's identity is constituted by its values. The presence of other subjects challenges my hierarchy of values (and thereby my identity), since they condition my values. I have to take into account how my values are integrated with the existence of other subjects with different values. Otherwise, I cannot interact in a world made up by a multitude of different subjects. I am continuously faced with the other subject and have to respond for my actions (i.e. values) to that subject, other than to myself. We might say that the presence of the other discloses the question of the legitimacy of my values. Indeed, a legitimate value is not a value prone to promote my own well-being at every cost, but a value that contains both a concern for my well-being and a concern for responsibility to other subjects. In this way, identity of the subject is constituted by the constellation of value, legitimacy, and responsibility. Subjectivity is, at its core, shaped by both itself and the existence of the other.

So, how does the individual subject organize its hierarchy of values in such a way that it satisfies the demands of legitimacy issued from both its own wellbeing and the responsibility for the well-

being of the other? Here Ricoeur structures his answer according to the basic scheme developed in the analyses of the transcendental synthesis of sensibility and reason.

The Person in between Character and Happiness

In the analyses of the transcendental synthesis, we saw that the perception of the subject is, due to its bodily constitution, linked to a certain perspective. This structure is now applied to the subject as a totality (FM 49/67). This totality is 'all the affective and practical aspects, all the attendant values and counter-values that attract or repel, all the obstacles, the ways and means, tools and instruments that make it practicable or impracticable and, in any event, difficult' (FM 47/65). Up until now, we have dealt with the subject as experiencing and interacting with the world and the other, but now Ricoeur gathers these different aspects (active as well as passive) of the subject together in a totality, i.e., in the concept of a person. This is the final move of the structural analyses of the subject. How does the subject become a person? As we shall see, this question is closely linked to the question of legitimacy because to be a person is to behave according to legitimate values.

Ricœur structures the concept of person on the transcendental scheme of perspective and meaning (FM 64/81). We must remember that the concept of perspective accounts for the finite capacities (body, receptivity) of subjectivity, whereas the concept of meaning expresses the infinite capacities (reason, spontaneity). These two capacities constitute the synthesis of conscious experience by means of transcendental imagination. Ricoeur now transfers this synthesis of finitude and infinity to the totality of the practical subject. The three concepts employed in this practical synthesis are: 1) Character (finite), 2) Happiness (infinite), and 3) Respect (mediation): '*All the aspects of "practical" finitude that can be understood on the basis of the transcendental notion of perspective may be summed up in the notion of character. All the aspects of "practical" infintude that can be understood on the basis of the notion of meaning may be summed up in the notion happiness. The "practical mediation that extends the mediation of the transcendental imagination, projected into the object, is the constitution of the person by means of respect*' (FM 49-50/67. Translation slightly modified).

The practical synthesis is, as the transcendental, analyzed by the concept of intentionality. Character and Happiness are not analyzed introspectively as inherent features of the subject, but understood through their mediation in the object, namely '*the constitution of the person by means of respect*' (FM 50/67. Translation slightly modified).

Earlier, we saw that the person is not a fact, but a problem and a task for the individual subject because the totality of the subject is a mediation between the heterogeneous aspects of subjectivity. Indeed, the disproportion between these aspects is present at every level (theoretic as practical) in the concept of subjectivity and makes possible the risk of failure. In the practical context, the disproportion is between character and happiness, and the issue at stake is the affirmation of the subject as person. Now, I shall briefly sketch the two concepts of character and happiness and then clarify how they are mediated in the person through the notion of respect.

The individual subject develops, through its continuous interactions with the world, a certain character: 'character is a limitation inherent in the mediating function of one's own body, the primal narrowness of my openness.' (FM 59/76). My character is a mixture of contingence and necessity in the sense that in my autonomy I am always dependent on someone or something (FM 63/80). I am born by two other persons and into a certain culture; I am endowed with at certain physical constitution; I am affected by events that are out of my control; I acquire certain habits through my existence in time. These facts form my existence, and they all condition my openness to the world. In other words, they, gathered in the concept of character, determine how I am affected by the world. In this way, '[m]y character is the primal orientation of my total field of motivation, and this field is my openness to humanity' (FM 63/81). My character is the expression of my 'practical finitude' because it influences how an object appears in the practical space of reason. The object stirs up a certain inclination or wake desire, which influence my motivation for a certain decision (FM 69). My decision, immediately, goes in a certain direction determined by my character. In consequence, my character 'is the individual who I am' (FN 368/345) and '[c]haracter is always my own way of thinking, not what I think.' (FN 370/347). This bears considerable effect on my values, because '[e]ach value is a universal which each individual stamps with his individual mark' (idem; FM 61/78). But fortunately, the subject is not only made up of its character. The character of the subject may configure the practical space of reason in that it determines initial inclinations and desires, but it is reason that has the last word (FM 66-7/84).

Reason is our 'practical infinitude' in the sense that it seeks to hold together the single decisions into a totality (FM 68/86). Reason is not satisfied by momentary pleasures, but seeks something else, namely the sum of all pleasures. Ricoeur defines this sum (as a faithful pupil of Aristotle) as the concept of happiness (FM 66/84). This, however, is a naïve conception of happiness (FM 82). Happiness is not a finite concept and can therefore not be satisfied by the sum of finite pleasures: 'Happiness is something entirely different; it is not a finite term […] Just as the world is the horizon

of *the thing*, happiness is the horizon from every point of view.' (FM 65/82). Happiness transcends the world of finite and perishable things and points to a demand of totality (FM 65-6/83). The totality expressed in the concept of happiness is 'the demand for a totality of *meaning*' (FM 66/84; FM 136/152). Happiness is not this or that happy state that appears and vanishes, but an idea that confers a meaning to my existence as a totality.

As we saw with the transcendental synthesis, meaning is conferred by reason in the way that the different objects become these particular things in this specific world by the work of reason. Reason demands a totality of meaning of the different stimuli in conscious experience. Thus, our conscious experience of the world is intentional in the sense that we understand the world by means of the meaning conferred by our reason. This is not the case only in conscious experience, but also in our practical interaction with the world. Reason demands that the totality of our actions and affections can be gathered in a meaningful whole, namely in happiness.

Therefore, when we decide on which value to follow in the practical space of reason, our decisions are not only induced from our immediate inclinations and desires (character), but also from a demand of totality issued from reason (happiness). The subject has, in its decisions, to mediate in the disproportion between the inclinations of its character and the demands of happiness. How, then, does the subject perform this mediation successfully?

Ricoeur now introduces a third concept, which is supposed to function as mediation in the way that imagination did in the transcendental synthesis. Character and happiness must cover the totality of the subject and their mediation has to be done in such a way that they form a synthesis, 'and that synthesis is found in the person.' (FM 69/86).

My character and my pursuit of happiness are to be mediated through the intention of becoming a person. 'The person is still a projected synthesis that seizes itself in the representation of a task, of an ideal of what the person should be.' (idem). My practical space of reason has to be configured according to this ideal so that the hierarchy of values comes to reflect the nature of personhood, namely an existence together with other persons. The values according to which I act in the world are shaped by my existence together with other persons, and I cannot disregard the happiness of the other person because of this coexistence. My happiness depends upon the others, because my existence is not constituted only by myself. Human existence is 'an existence that one apprehends, or, to be more precise, a presence with which one enters into relations of mutual understanding, exchange, work, sociality.' (FM 71/88). To enter into the society of persons, my hierarchy of values has to be ordered according to the notion of humanity. I have to recognize the humanity in the other

in order to interact with her. We must understand the other person in 'his "belonging" (as a member or as a leader) to a practical and ethical totality of persons. Outside of this, one is no longer a person. One's existence can only be a value-existence' (KH 200/250). Only by recognizing this can I legitimize my values to myself and to the other. Recognizing the humanity of the other person is not arbitrary, since the happiness of my own existence depends on it. Furthermore, humanity is not only a theoretical issue, but an acting in recognizing that '[h]umanity is a way of treating human beings, you as well as me. It is neither you nor I, it is the practical ideal of the "Self" in you as well as in me' (FM 72/89. Translation slightly modified).

I confer humanity to my practical space of reason by treating the other person with respect. Ricoeur employs the Kantian notion of respect in order to answer the question about the legitimacy of the person's values. In deciding on what value to follow, I have to respect the humanity that I recognize in the persons in front of me. My practical space of reason is not only rooted in my own being, but at the same time in the being of all other persons. And that which we, human beings, have in common in spite of all our 'existential differences' (perspectives and characters) is exactly our humanity. In respecting the other person as a person, never as a means, I myself confirm my personhood, my humanity. The Kantian notion of respect functions as the mediation between the two fundamental aspects of the practical space of reason, character, and happiness. It operates in the same way as imagination did in the transcendental synthesis in that it mediates in the disproportion at the core of the subject (FM 73/90).

We now see that my practical space of reason is not the experience of an individual subject, but a space configured by the coexistence with other persons. It is a 'practical and ethical totality' (KH 200/250) to which I appertain in virtue of my humanity. My values and actions cannot be separated from those of the other. They involve and affect the being of the other person, like her values and actions involve and affect my being.

The subject, however, has a hard time living up to the ideal of the person. It often fails to treat the other persons with respect and, often, it configures the values in the practical space of reason according to its own self. Kant blames sensibility for this degeneration of personhood. Our inclinations, desires, feelings and emotions distort the clear imperatives of reason. The respect for the happiness of the other person is suffocated by the person's sensibility (bodily perspective and character), and, therefore, sensibility must be extirpated from the practical space of reason. The person must obey reason and avoid the deceptive sensibility (FM 75/92). Ricoeur, though, opposes such a 'ethic dualism' (FM 79/95), and the consequently pessimist view on human sensibility as

'the radical evil'. Reason and sensibility are intertwined even in the most basic core of subjectivity (which Kant himself was very conscious of), and therefore we have to 'explore the "practical" disproportion that is more primordial than the ethical duality, and *to uncover a principle of limitation that would not already be radical evil*' (idem. Translation slightly modified).

Ricoeur concludes his structural (stripped) notion of subjectivity with an account of how the failure to live up to the practical ideal of the person is rooted in a fundamental disproportion of reason (mind) and sensibility (body) in the heart of the subject. The fragility of the person derives from this original disproportion in subjectivity. The structural analyses have purposely disregarded the affective aspect of subjectivity concentrating on how the disproportion acted '*on* the objectivity of the thing, *on* the humanity of the person' (FM 81/97-8). In a final step, Ricoeur turns to the question of how this disproportion is experienced by the subject as a self (idem). To explore this self-experience, Ricoeur turns to the question of affectivity.

Affectivity: The Conflict of Feelings

In this concluding section, I will look on three notions that together unveil the ontological dimension of the foregoing analyses. First, the notions of disproportion and fallibility that tell something about how the subject experiences and understands the world, the other and itself and thereby determines its decisions and actions. Then, I analyze the complex notion of affectivity in order to clarify its function in subjectivity and what it reveals about the ontology of the human subject. In a way of conclusion, I prepare the next chapter by explaining how the stripped notion of subjectivity naturally leads to questions that Ricoeur deals with and elaborates further in his later writings.

<u>Disproportion and Fallibility</u>

Subjectivity is basically structured by body and reason. This implies, as we have seen, that the subject is always bound to a certain point of view that limits its openness to the world. Nevertheless, the subject can, through reason, transcend this perspective and try to understand the meaning of the limited perception of the world. This basic transcendental structure clarifies a fundamental condition of subjectivity: the subject has to mediate between the two heterogeneous capacities of its nature. The relation between the two capacities is, however, marked by a disproportion. The concept of disproportion was introduced with the transcendental synthesis (FM 37/55). In the synthesis, transcendental imagination plays a key role in that it mediates between

body and mind in the generation of conscious experience. Experience is basically a result of how the subject judges that which affects it. This is central point in Ricoeur's theory. The subject is an embodied mind and not a reasonable body. Although Ricoeur never underestimates the importance of the body in the constitution of subjectivity, he has no doubt about the primacy of reason. The subject is intelligible only from the perspective of its capacity to reason (i.e. its capacity to judge between possible options and values). He repeatedly stresses this point in his early works (FN 8, 37, 57, 342; FM 62, 81). Thus, there exists a disproportion in the subject. The mind dominates the body, but is nevertheless bound to it and limited by it. Together with the concept of disproportion, Ricoeur introduces another critical concept for the condition of subjectivity, namely the concept of fallibility: 'We are looking for fallibility in the disproportion' (FM 1/22).

These two concepts, disproportion and fallibility, are uncovered by the structures analyzed above. In the transcendental generation of conscious experience, the subject is not in an immediate relation to the world or to itself. The experience of the world and the understanding of the self as a separate and autonomous part of that world always involve an activity on behalf of the subject. In conscious experience, the subject have to judge about the stimuli affecting it through the bodily senses. In the forming of personal identity, the subject must choose the values by means of which it interacts with the other subjects. In both cases, the subject is exposed to the risk of making wrong judgments. The risk is always highly present, because '[t]he bond with the body, even though indivisible, is polemic and dramatic' (FN 227/212; cf. FN 485/456). This tense and difficult situation is a result of the disproportion of sensibility and reason. On one side, reason is the capacity by which the subject dissociates itself from the workings of the world and, thereby, makes room for freedom. On the other, the body is that which, inescapably, binds the subject to the workings of world.

Although reason is that which enables the subject to become a person (through decisions, i.e. actions), it cannot disregard its bodily constitution. In every decision the subject has to pay attention to both its (potentially) unlimited capacity to reason and the limits imposed by body: 'It is to the extent to which the entire world is a vast extension of our body as pure fact that it is itself the terminus of our consent' (FN 343/321; cf. FM 21-22/40). The body proper affects and restricts the subject's access to the world. And the subject cannot ignore or, in any way, escape this 'pure fact'. The body is both access (registration of stimuli from the surroundings) and limit (the access is always determined by a specific perspective) to the world. The subject must use reason in order to give meaning to the world accessed by the body. A threefold cord, 'the notions of perspective, meaning, and synthesis' make up 'the melodic germ of all the subsequent developments' (FM

49/66). The risk of failure in this synthesis is always present. This fragility affects the most trivial, and often unconscious, decisions (such as classifying objects as certain things), but becomes much more vulnerable when the subject is in front of the other subject. The subject is a self, this particular subject with its own desires, fears and ideas, in short, with its own character, and yet it is situated in a world composed by other subjects (with different characters).

Hence, we see that disproportion and fallibility have been following us at the different levels of the analysis. Now, in a final analysis, Ricoeur employs these structural features of subjectivity (infinite verb vs. finite perspective and happiness vs. character) in order to find out how the subject itself experiences the disproportion and fallibility rooted deep in its constitution. He turns to what he claims to be the core of the disproportion, namely the affective dimension of subjectivity (FM 91/108). With the affective dimension, he hopes to immerse the structural analyses in the pathetic (i.e. suffered, affected) nature of human existence. Human existence is pathetic, because our problematic relation to the world, the other and ourselves is *felt* before it is *understood* (FM 81/97).

The philosophy of feeling (idem) interiorizes the structural features in that it uncovers the disproportion in the heart of the subject and shows how fallibility is rooted in the fragile constitution of subjectivity: 'That is to say, the constitutional weakness that makes evil possible' (FM Xliii/11. Translation slightly modified).

The Nature of Feeling

Affectivity is the dimension of subjectivity revealed by feeling. The existence of the human subject is first of all characterized by feeling, 'the affection through which I feel myself existing' (FM 131/147).

The values that form the practical space of reason are not explained by intentionality. Intentionality may give the names 'good' and 'bad', but does not explain *why* our experience is qualified in this way (FM 90/106). Feelings, on the contrary, reveal *why* an object or a situation is good or bad. Reason organize our hierarchy of values according to the basic structures of the subject in relation to the world and the other subjects, but feelings are what make this matter to me in the quality of the subject that I am: 'feeling interiorizes reason and shows me that reason is my reason, for through it I appropriate reason for myself […] In short, feeling reveals the identity of existence and reason: it personalizes reason.' (FM 102/118; cf. LS 260). My first-person experience is permeated and qualified by feelings. But what are feelings? The term feeling covers a multitude of particular

functions: ‘affective regulations, disturbing emotions, affective states, vague intuitions, passions, etc.’ (FM 83/99).

The function of feelings is best explained in relation to reason. We understand the world by means of the transcendental synthesis in that we categorize the information received through the senses as objects and position these objects in relation to ourselves. We understand by the opposition between subject and object. Now, fellings inverses this objective form of understanding because ‘[l]e sentiment est la manifestation sentie d’une relation au monde plus profonde que celle de la représentation qui institue la polarité du sujet et l’objet’ (LS 253; cf. FM 85/101); in short, ‘il relie d’abord ce que la connaissance scinde’ (LS 264). Feelings disclose my existence as being already situated in the world that I attempt to understand. I am in the world before I understand it.

Still, feelings and reason are not two separate functions in subjectivity, but work together as a unity. The structural analyses disregarded the feeling dimension in order to sort out the formal structures of subjectivity, but in the situated existence of the subject the capacities are generated through and by one another (LS 252). Feeling discloses the intention of reason, and reason clarifies the intentionality of feeling.

When I consider a certain action, experience a situation or meet another person, the feelings reveal the importance of these phenomena to me. I may feel that the person is lovable, the action repugnant or the situation embarrassing. Feeling exposes my interiority by manifesting ‘pre- and hyper-connections with the beings of the world’ (FM 86/102). We can call this relation the personal relation to the world. It is personal because feeling is what qualifies experiences as *my* experiences. It is *my* love, *my* repugnance, and *my* embarrassment. On the other hand, feeling is always a feeling of something. The qualitative dimension of feeling is always bound to an object (physical or non-physical) (LS 252). There is intentionality in feeling that points to something other than just *my* feeling: ‘Our “affections” are read on the world they develop, which reflects their kind and nuances’ (FM 84/100). The intrinsic interplay of feeling (interiority) and reason (intentionality) in affectivity discloses a critical feature of human subjectivity: the coexistence of *self* and *the other than self* in subjectivity itself. The subject is, in its inmost core, affected both by itself and by that which is different from it. Feelings are important because of this duality of affectivity. They are, at the same time, generated by the world, the other and the subject itself. They reveal a complex dialectic between selfhood and alterity, the voluntary and the involuntary, activity and the passivity, in the core of the subject. The existing subject is not a free-floating being with an unlimited capacity to choose; on the contrary, it is a being situated in and bound to a context, a world and other subjects

that determine and shape its existence. The affective dimension is exactly what reveals disclose our 'being already in…, through that primodial *inesse*.' (FM 103/119). In the analysis of the practical aspect of the subject, we saw that human action is determined by the character of the acting subject (habits, physical constitution, education, culture etc.). The subject is inclined to satisfy the values that are shaped by its character. However, we also noticed that the presence of other subjects call into question the legitimacy of those values. The subject is intertwined with the other in its existence in the world. Action in this coexistence of different subjects takes place in what I chose to call the practical space of reason. The practical space of reason is the world experienced as a coexistence with other subjects considered as persons. It is the world pervaded by values generated and shaped by the self, the other, and the world. My feelings reveal to me that I 'am-already-in' this practical space of reason. My existence is an existence in a world constituted by heterogeneous and discordant values.

The complex nature of values is revealed by the affective dimension, since this dimension of subjectivity, as opposed to the transcendental and the practical, expresses 'living consciousness' (FM 108); an active, living consciousness is related to the world and itself 'par tous ces fils secrets, »tendus« entre nous et les êtres, que nous nommons précisément les »tendances«' (LS 253). These tendencies are the origin of our values and motivations. They disclose the *why* of the predicates 'good' and 'bad'. A tendency holds this power to qualify since it, though feelings, solicits a subject in direction of a certain action. It stresses the involuntary aspect of values by linking them to basic needs such as eating, drinking, sleeping and reproduction. However, it is difficult to speak about tendencies in other animals than humans, since feelings in animals seem to coincide with their instinctual behavior. In this, human behavior is different, which becomes clear when we look at the feelings involved in human tendencies. The human subject is affected by an obscure confusion of tendencies due to the disproportion inherent in its being. We belong to both biological nature and the sphere of persons, or culture (as Ricoeur calls it). Our tendencies drive us toward both these dimensions which often situate us in a state of conflict or tension. In fact, the function of feelings has to be understood in relation to this tension: 'Feeling points out how far along we are toward the resolution of tensions. Its modalities and its felt nuances mark the phases of action launched by a certain disequilibrium seeking a new equilibrium. Feeling is thus a function of the recovery of the living creature's equilibrium. To understand its role in this process is to understand feeling.' (FM 99/115. Translation slightly modified). The job of feelings is to mediate in the conflict between sensibility and reason. The notion of conflict has been following us from the outset. The other kinds

of conflicts, though, have been dealt with as formal conflicts with regard to subject's objectal relation to the world and the other (*the thing* in the transcendental synthesis and *humanity* in the practical). Now Ricoeur wants to find the original conflict in the nature of the subject itself. And he finds this conflict in the feelings, because they disclose 'the inner conflict of human desire' (FM 92/108)

All human conflicts originate in what Ricoeur calls the 'affective node' of subjectivity (idem), which is the radical and dramatic, i.e., the felt, manifestation of the original duality of human subjectivity; the disproportion felt as conflict (FM 106/123). To approach the intimate affective core of subjectivity, Ricoeur develops the Platonic treatment of the relationship between the heart [thymós], the sensous desires [epithymía], and the spiritual desires [éros]. Vital and spiritual desires keep the subject in a continuing tension between 'pleasure' and 'happiness' felt due to the fact that our being is situated in the world as both a biological and spiritual being. Our desires drive us towards a complex satisfaction of values generated by both of these aspects of subjectivity. However, the subject is not divided into two separate parts, a pure reason and an obscure sensibility, but acts as a unity by means of a mediation of vitality and spirituality. And this mediation takes place in what Ricoeur calls the heart. The heart is, at the same time, organ and symbol for the feelings involved in existence (FM 104/120). The notion of the heart secures that there is no ontological difference between (organic) pleasure and (spiritual) happiness. In this way, Ricoeur rejects the essence of Kantian and other philosophical anthropologies (the Thomist and the Cartesian), namely that the subject has to obey only the rational part of its nature in order to become a person (FM 77-8/92-3). For Ricoeur, humanity of the subject, personhood, is not constituted only by reason; on the contrary, it makes no sense to speak of a pure reason, since human reason is always rooted in sensibility, just as sensibility is always shaped by reason. As we have seen earlier, the subject cannot be understood as a static being made up of different parts, but only through its unity in actions. The subject is not given before its actions. Therefore, the complexity of human desires has to be understood through the behavior that defines the subject: 'Une tendance, c'est à la fois la direction objective d'une conduite et la visée d'un sentiment; aussi le sentiment n'est-il rien d'autre que cette direction meme de la conduite en tant que sentie; la manifestation ressentie de ce »vers quoi« s'approche, »loin de quoi« s'éloigne, »contre quoi« lutte notre desir' (LS 253). It is the felt relation to that which is not the subject (the world and the other) that is expressed in heart. Our conduct is always a product of the manifold of tendencies, vital as spiritual, felt in the heart as desires. Our desires can thus be both vital desires (pleasure) and spiritual desires (happiness), but it

is important to emphasize that both pleasure and happiness are felt as desires in the subject (FN 124-5/119). We must go beyond the traditional conception of desire as pathological aberration and delirium and seek out the authentic desire at the core of all desires, 'the "quest" of humanity behind the passional "pursuit", the quest that is no longer mad and in bondage but constitutive of human praxis and the human Self' (FM 111/127). This authentic desire is a seeking for happiness, understood as the totality of my actions as a meaningful whole. Whereas the vital desires seek immediate satisfaction of need (i.e. biological ones such as eating, sleeping, and reproduction), the spiritual desires aim at a satisfaction of my existence as a totality (i.e. my coexistence with other subject in friendship and society, my convictions, my dreams, etc.). The heart of the subject is in a continuous tension between these heterogeneous desires, which is experienced in feelings. Ricoeur separates the manifold of feelings into two categories, the schematized and the atmospheric feelings.

The schematized feelings are those involved in our interpersonal relation to the other subjects. By schematization, Ricoeur intends the categorization of the feelings involved in the constitution of our practical engagement with the world, i.e., the feelings involved in our decisions, values, and actions that constitute our coexistence with the other in the practical space of reason: 'we must specify and articulate the relationship of the Self to another Self by means of the objectivity that is built on the themes of having, power, and worth.' (FM 113/129). Human feelings can only be understood in the intentional relation to the other person through the dimensions of politics, economics, and culture (idem). Possession, power and value expressed in these 'new' dimensions awake a multitude of feelings at the heart of the subject. They disclose the feelings involved in our concrete existence with the other Self.

Not all feelings, though, are expressed by an intentional relation. We experience feelings that do not seem to refer to any distinct object. They are more like an atmosphere or a tonality accompanying our being situated in the world without relating to anything in particular. Thus, Ricoeur calls these feelings atmospheric. They are the formless background upon which we conduct our lives. Their origin is the 'being-already-there' of our existence, and they are, in this way, the felt expression of the involuntary or passive aspect of human subjectivity. They are experienced as moods (being in a good mood, a bad mood, lightness, heaviness, welfare, uneasiness, happiness, joy etc.), gathered as the totality of our sentiments in 'the fundamental feeling' (FM 105/121). It is in this fundamental feeling that the schematized feelings originate, from which they develop into distinct feelings with an intentional content, and into which these dissolve themselves again (idem).

This fundamental atmosphere of the heart is the internalization of the subject's relation to the world as a totality, '[b]ut by interiorizing all the connections of the self to the world, feeling gives rise to a new cleavage, of the self from the self. […] It stretches the self between two fundamental affective projects, that of the organic life that reaches its term in the instantaneous perfection of pleasure, and that of the spiritual life that aspires to tonality, to the perfection of happiness.'(FM 131-2/148).

This 'cleavage' in the subject reveals the ontological status of the subject in the form of conflict. We have met the concept of conflict several times in the previous analyses. Now it is time to confront this conflict in terms of an ontology of the subject.

Outline for an Ontology

Being this particular subject situated in a certain world is always characterized by feelings. Some feelings (the schematized) are explicitly related to our interaction with the world, others are mere tonalities or atmospheres that accompany our existence as the background mood from which we embark upon our doings and in which we sustain our sufferings. Ricoeur said that feelings, basically, function as affective regulations that aim at 'the re-equilibration of the living', and to understand their function in this process is to understand the feeling itself. Then, how do feelings regulate the being of the subject? To answer this question, we have to return to the notion of disproportion.

The disproportion of reason and sensibility extends from the most basic structure of subjectivity (transcendental synthesis of conscious experience) through the ideal of the person in action (the practical space of reason) to the intimate feeling of being a self (affectivity). The intimate feeling of being a self is experienced as an atmosphere, a fundamental feeling, rooted in two heterogeneous aspects of the subject, the vital and the spiritual, that each awakes certain desires. On one side, desires that impel towards satisfaction of immediate pleasure, and on the other, desires that aim at the existence as a totality of meaning or, in other words, happiness. These two kinds of desire are in tension in the heart of the subject, which is expressed in the fundamental background feeling, the mood, in which the subject exists. This can be a feeling of anguish, sadness, joy, satisfaction, pleasure, and so on (FM 106/122). Our felt relation to the world, the schematized feelings, springs from this background, motivate our actions by confering our values their embodied and personal meaning. It is the heart that mediates between the two heterogeneous desires, and in the heart we find the original conflict that characterizes human subjectivity: the separation of the self from the self. This non-coincidence is felt as a conflict in subject: 'It seems, then, that *conflict* is a function

of man's most *primordial* constitution; the object is synthesis; the self is conflict' (FM 132/148). Our feelings regulate this conflict in the sense that they personalize the values and transform the objective structures in the world into a practical space of reason, wherein we see ourselves and engage with the world and the other through a hierarchy of values.

The inmost intimacy of the subject, the heart, is in continuous unrest because of the conflict between the two heterogeneous aspects of its being, body (vital desires) and reason (spiritual desires), and feelings are what reveal 'this non-coincidence of self to self' (FM 141/157). The objective disproportion between verb and perspective, character and happiness that we analyzed above, now finds its 'ontological "locus"' (FM 134/150) in the complex constitution of the subject. It is through affectivity that we find the ontological nature of the human being as an intermediary being (FM 108), constituted by both finitude and infinitude (FM 134/150).

But even though conflict is constitutional of human subjectivity, Ricoeur has no doubt as to the relation between the two aspects of our nature. We have to understand the subject from its power to affirm, its infinitude, which in our foregoing analysis has been expressed in the verb, the idea of happiness and the sensible happiness of the heart (spiritual desire), 'Feeling alone, through its pole of infinitude, assures me that I can "*continue my existence in*" the openness of thinking and of acting' (FM 137/153). Our capacity to think and act springs from the originating affirmation, which is 'the effort to exist' (FM 137/154) that characterizes our being. However, 'the originating affirmation *becomes* man only by going through the existential *negation* that we called perspective, character, and vital feeling' (FM 137/153). This 'existential negation' is the fact that my being is contingent or out of my hands, so to say: 'Existence is discovered to be *only* existence, *default of being-through-self*' (FM 139/155).

Thus, Ricoeur defines the human subject as a being in conflict between an originating affirmation (a will to exist) and an 'existential negation' (perspective, character, vital desires). The self is not a being immediately given to itself, but *becomes* itself through the conflict intrinsic in its nature (FM 141/157). The self is the fragile synthesis between these two aspects of its selfhood. On the one hand, the self is rooted in the vital will-to-live, which determines its desires and actions in the direction of self-preservation. I am only a self as long as I breathe. On the other, it is a person intertwined with other persons in virtue of its humanity, and it can only become a self by being and acting by the desire to become a person through and in this humanity. The analysis of practical synthesis clarified that we have to act in accordance with this idea of humanity, but it is only through affectivity that we understand that we *are* this humanity; that our being is an existence in

the idea of humanity because 'then reason is no longer an other: I am it, you are it, because we are what it is' (FM 137/153). Our feelings reveal this duplicity in our nature as a conflict between our desire to be what we are, the continuing existence of this particular subject, this self, and our desire to be a person among persons, a self understood as more than itself, as a part of humanity: 'The community is my good because it leads towards making me whole within a "we" where the lacuna of my being would be filled' (FN 128/122). The subject cannot disregard any of these aspects. It is both this particular subject *and* a part of humanity. Human fragility is due to this ontological status of the subject as a mixed nature (FM 156). And this peculiar nature determines the subject's actions because '[h]uman freedom is a dependent independence, a receptive initiative' (PW 228/79).

At this point, however, it is important not to misread the notion of existential negation. It is not a flaw in or degeneration of the subject. It is an inexorable ontological trait of human subjectivity better understood as an 'existential difference' (FM 135/152) that qualifies the subject as the this particular subject, a self. It is the contingency or alterity that characterizes the existence of the human subject. It is that by which we differ from each other, but at the same time also that which, together with the originating affirmation, enables 'the understanding of language, the communication of culture, and the communion of persons' (FM 138/154). The ontology of our being human is rooted in a continuous tension between selfhood and alterity felt as a conflict inherent in the self: 'the Self is never certain' (FM 126/142). The self is in restless search of restoring itself as a self.

The notion of existential negation or difference is critical for the further development of Ricoeur's theory of subjectivity because the subject only finds itself thorough this difference. Since the being of the subject is marked by a non-coincidence between the self and the self, the subject must turn towards that which can restore its being in existence, namely the existence together with other subjects in time, language, and society. We saw how the fundamental feeling of being, the atmosphere of existence, transforms into schematized feelings through the notions of 'having, power, worth'. These schematized feelings are characterized by their object-directedness in the sense that their meaning is produced in the encounter between the subject and alterity. The subject is driven towards the world and the other in order 'to fill the lacuna' in its own being.

Therefore, Ricoeur makes a change in methodology in order to approach this detour over the meaning of the encounter with alterity. He turns to the language, culture, and temporal being of the subject in its coexistence with the other subjects to discover the restoration of selfhood.

This detour over language, time, and coexistence is the object of the following analyses.

Chapter Two

The Redressed Notion of Subjectivity

This chapter follows closely on the heels of the foregoing. It continues to reformulate Ricoeur's theory of subjectivity as strict as possible. Something surprisingly takes place, however, in Ricoeur's theoretical works after FM (1960). Ricoeur still engages in the same questions: how does the human subject experience the world? How does it act? How is it affected by its existence in the world? But whereas his early works dealt with these questions by means of a reflective-phenomenological method (intentionality, synthesis, and originating affirmation), he now turns to hermeneutics. But this shift in method is not arbitrary.

The stripped notion of subjectivity, uncovered by structural analyses, made it clear that the fragility of the subject stems from a conflict inherent in subjectivity, namely the conflict between sensibility and reason. We followed this conflict in three steps, from conscious experience over the basic structures of action in the configuration of the practical space of reason to the felt conflict in the heart of the subject. The conflict grew more dramatic until it culminated in the felt conflict between selfhood and alterity in the core of the subject (vital and spiritual desires).

The following analyses will concentrate on the question of human identity in relation to how the subject acts in the practical space of reason. We saw that the dimension of affectivity confers the practical space of reason the pathetic dimension of actual existence. It interiorizes the intentional attitudes involved in action and, thereby, gives experience and actions their meaning (i.e. value) for the subject as an individual self. The subject is driven by an originating affirmation of existence (expressed in the verb, practical totality and happiness), but this affirmation is always conditioned by an existential negation or difference (bodily perspective and character), felt as the vital desire to exist on the terms of my body and character. The subject has to mediate between these two aspects of its nature in order to become a subject inscribed in humanity. And only through humanity can it satisfy the desire of happiness elicited by the originating affirmation.

At this point, the importance of the humanity of the subject appears more as a hypothesis than an argument. The previous analyses have showed that the happiness of the individual subject is necessarily related to the other, but have only circumscribed the conflict of selfhood and alterity in the heart of the subject. To develop and enrich the concept of humanity in relation to happiness and argue for a strict connection with the conflict between selfhood and alterity, we have to turn from the structural to an ontological analysis of subjectivity.

Once clarified the structures of experience, action and affectivity, we must place the subject in the world again. One of the central discoveries of the structural analyses was that the subject is not immediately a self, but marked by non-coincidence. The subject has to *become* a self by mediating the structural conflict in its nature. This conflict may be rooted in the inmost core of the subject, the heart, but, through the intentionality of the feelings, it expresses itself in the subject's interaction with the world. This interaction takes place in the practical space of reason, the world experienced as a configuration of values, where the subject finds itself lost among the objects of the world and their heterogeneous values. It has to *reappropriate* itself through the interaction with these objects: 'Appropriation signifies that the initial situation from which reflection proceeds is "forgetfulness." I am lost, "led astray" among objects and separated from the center of my existence, just as I am separated from others and as an enemy is separated from all men. Whatever the secret of this "diaspora", of this separation, it signifies that I do not at first possess what I am." (FP 45/53; see also A 192). To find oneself through reappropriation is another expression of the mediation of the conflict between selfhood and alterity, which is now approached as a concrete problem of existence. How does the subject affirm itself in the historical coexistence with alterity (objects and, more importantly, other persons)? Only through alterity can the subject affirm itself as a self. The unrest and uncertainty that characterizes the fundamental, ontological feeling of the subject is rooted in the non-coincidence of subjectivity. We noticed in the beginning that Ricoeur characterizes his philosophy as a restoration of the originating affirmation, the basic desire-to-live; now, he specifies that this restoration can only take place in a reappropriation of the self situated in a world which is constituted by selfhood and alterity. The subject may be an immediate, originating affirmation, but this affirmation is lost in the coexistence with alterity.

The subject finds itself situated in a world in which it has not placed itself, and it is in this world, in interacting with alterity, that it must reaffirm itself as an individual subject; a self that is an I am…: 'the self [le moi] must be lost in order to find the "I" [le je]' (EH 19/24). The existing subject is in continuous struggle with the question, *who am I?* The reappropriation of the self must measure itself with the insecurity stirred up by the alterity that constitutes its being: 'The hermeneutic of the *I am* can alone include both the apodictic certainty of the Cartesian *I think* and the uncertainties, even the lies and the illusions, of the self, of immediate consciousness. It alone can yoke, side by side, the serene assertion *I am* and the poignant doubt *Who am I*?' (QS 259/262).

In this chapter, I shall attempt to clarify this 'hermeneutic of the I am' in close connection to the structural analysis. It is an attempt to articulate what Ricoeur names 'an ontology of selfhood in

terms of actuality and potentiality' (OSA 308/357); in other words, what the subject is and what it can be. The chapter is, as the previous, divided into four sections. First I begin with a clarification of the methodological concepts of hermeneutics and narrative identity. This section centers on the works *Freud and Philosophy* and *Time and Narrative 3,* but draws substantially on minor articles regarding the hermeneutical turn. I shall deal thoroughly with the concept of hermeneutics in order to explain the reasons for the turn to hermeneutics and only briefly introduce the concept of narrative identity, since this concept is to be developed and explained further in the remaining sections of the chapter. The following three sections inverses the scheme adopted in the first chapter (from experience over action to affectivity) in that it begins with affectivity, then proceeds to action and ends with ethical experience. In this way, I hope to show how the structural analysis is a necessary precondition for understanding the scope and meaning of Ricoeur's theory of subjectivity. The main references for the analyses in these sections are *Oneself as Another* and *The Course of Recognition*.

Basic Methodological Concepts

The subject has no immediate knowledge of itself. We learned this from the stripped notion of subjectivity. Nonetheless, it has a strange, almost paradoxical, ring to it. If the subject were supposed to know anything at all, then this something should be itself, or at least its capacity to assure itself of itself through thinking. This was the assertion of the Cartesian Cogito and the fundament upon which the modern self was explored and developed, but something happened with 'the school of suspicion' whose most prominent masters were Marx, Nietzsche, and Freud (FP 32-3/40-1). The idea of self-knowledge changed from an assertion to a question. Consciousness is not immediately given, but a 'task' (FP 44/51). We can no longer content ourselves with the idea of an invincible first truth about the self as an immediate intellectual intuition, a psychological evidence or a mythical vision (idem). Such kind of knowledge is doomed to remain abstract and empty, or even worse, it may be used to lead us astray and into illusions about ourselves. The methodological lesson to be drawn from 'the school of suspicion' is that the subject may posit itself as a Cogito, '[b]ut it is a wounded [blessé] Cogito that results from this adventure – a Cogito that posits itself but does not possess itself; a Cogito that sees its original truth only in and through the avowal of inadequacy, illusion, and lying of actual consciousness' (FP 439/425). Immediate self-knowledge can no longer be used as a firm philosophical fundament, since the subject itself has to be included as part of the philosophical question and not as its presupposition. To accomplish the 'task' of

coming to know itself, the subject must refrain from the pretension of immediacy and understand itself as 'the watchful ego, attentive to its own presence, anxious about itself and attached to itself' (FP 54/62).

The pretension of immediacy was already weakened, if not expelled, in the previous chapter, since the basic unity of the subject was shown to be a constant mediation between the two heterogeneous aspects of subjectivity, namely sensibility and reason. This mediation is present at every level of subjectivity, even in a basic function such as conscious experience; and, as the analyses turned to subjectivity as selfhood, we saw that this mediation became a conflict between selfhood and alterity. The subject is primordially characterized by the capacity to affirm itself, to posit itself, to signify and to act, but this capacity is restricted by the bodily perspective and the character. The subject does not become a self but in the mediation of these two aspects of its nature. All possible knowledge has to pass through the 'existential difference' that differentiates one subject from the other. It is through my body that I enter the world as a space of meaning: 'A meaning that exists is a meaning caught up within a body, a meaningful behavior' (FP 382/372).

The subject must therefore find itself through its bodily existence, i.e., it has to understand itself as being situated in a world of objects that affect it through the body. We saw earlier that the subject becomes a subject only through action. Action, though, is not a result of a pure choice, but an interaction with the objects (and persons) in the world. The objects affect my action, just as my action affects the objects. The answer to the question 'who am I?' is to be sought in an interpretation of the subject's interaction with the world. It is not the subject who gives meaning to the object; on the contrary, it finds itself situated in a world of meaning that shatters its pretension to posit itself as a self: 'it is this inversion of thought which now addresses itself to me and makes me a subject that is spoken to.' (FP 31/39). The objects reveal my situation as being an individual subject, a self, in a world already loaded with meaning. However, this revelation of my existence as a self is not immediate, but 'remains an interpreted existence' (EH 23/27).

Hermeneutical Phenomenology

In what sense is existence always an interpreted existence? Before answering this question, we have to look more closely on the relation between subjectivity and language.

The previous chapter emphasized mediation as a critical feature of subjectivity: the subject is constituted by a mediation between the two heterogeneous aspects of its nature. The verb is that which confers a meaning to our limited perception. Without our capacity to designate, the objects in

the world would remain shattered perspectives and fragmentary information passing through our senses. This, however, is not the case. Our perception of the objects is characterized by meaning. I do not perceive an object as a multitude of different perceptions, for example, redness, solidness, smoothness, lightness, bent in different angles, but as a totality of meaning, a chair. This is what is meant by the intentional character of subjective experience: the object is already given to the subject as a meaningful whole, a thing, in the act of perception (FP 380/371). In the case of basic perception such as classifying objects as specific things, the mediation takes place without conscious attention on part of the subject (except in the cases where the classification goes wrong). This is complicated drastically when it comes to the question of the identity of the subject in interacting with the world. Perception is meaningful because it is rooted in the existence of the subject; and to understand the existence of the subject, we have 'to question back from the uttered meaning to the meaning in operation,' and understand why '[m]an is language' (FP 383/374); or put in another manner, '[i]t must be rediscovered with Hegel that language is the being-there of the mind' (idem).

The existence of the subject is characterized by language, since the subject interacts with the world *through* language. Perception finds its meaning in language. Hence, Ricoeur can say that the analysis of language is an extension of that of perception (idem). Our existence is not completed in the immediacy of perception, but protracts itself to that which is not present in perception, namely the past and the future. The human capacity to designate absent objects is what makes subjectivity an existence in time. It is through language that I can retain what has been and imagine what will be. Moreover, it is through language that the coexistence with other subjects is articulated and formed. I communicate and interact with others through language. However, Ricoeur is well aware of the complexity of the assertion that man is language: 'Language is no more the foundation than it is an object; it is mediation; it is the *medium*, the "milieu," in which and through which the subject posits himself and the world shows itself' (QS 250/252). He does not say that man is *only* language or that language is the sole medium through which we can understand human subjectivity. He merely stresses that man cannot be understood in isolation from language.

We should now be able to answer the question of the interpreted nature of human existence. The rejection of immediacy suggests that subject has to approach itself by a detour over the objects that constitute its being in existence, and since these objects reveal mainly themselves through language, a reflective recovery of subjectivity must necessarily be hermeneutic in nature.

We have seen that the subject, in its inmost core, is characterized by an originating affirmation, a basic desire-to-live, and that this affirmation is conditioned by an 'existential difference' (body, character and context). Now, the 'existential difference' is what makes the subject a human subject in that it situates the subject in a world whose meaning is derived not only from me, but also from the world and other subjects as well (subjects that are existentially different from me). I find myself embedded in a world of heterogeneous meanings and values, a practical space of reason, and can therefore only come to know myself in relation to that world of meanings and values. Meaning and value are not produced by me, but I interact with and am affected by them. And language is the medium through which meaning and value are revealed to me. I interpret my existence through the signs of meaning generated in the coexistence with the other subjects in the world. The meaning of my existence is not immediately given, but has to be recovered in an interpretation of my being-in-the-world: 'by passing through a hermeneutics, reflective philosophy emerges from abstraction; the affirmation of being, the desire and effort of existing which constitute me, find in the interpretation of signs the long road of awareness […] The appropriation of my desire to exist is impossible by the short way of consciousness; only the long path of interpretation of signs is open. Such is my working hypothesis in philosophy. I call it *concrete reflection*, that is, *the cogito mediated by the entire universe of signs*.' (QS 257-8/260).

We now see that the aim of hermeneutics is to situate the subject in the world. In the early works, Ricoeur clarified the basic structures of subjectivity by means of a reflective-phenomenological method; now this approach is complemented by hermeneutics in order to sort out the conflict of selfhood and alterity in the actual existence of the subject. Ricoeur still grounds his analyses in a phenomenological method, but it now becomes a hermeneutical phenomenology in that it develops the abstract notion of subjectivity into a lived subjectivity in a concrete world. In doing so, he attempts to answer the question about the identity of the subject in *existence*. The hermeneutical method approaches the notion of conflict by means of the concept of interpretation. Because of the non-coincidence of the self, identity is possible only through an interpretation of the subject's interaction with world and the other subjects. The subject must reaffirm itself in an interpretation of the coexistence with alterity; or '[e]ven better, it could be said that what is one's own and what is foreign are polarly constituted in the *same interpretation*' (PH 128/80).

The notion of the symbol might clarify the nature of the hermeneutical method. Symbols and metaphors have a prominent position in Ricoeur's philosophy. Here, I shall disregard the question

of metaphors and deal only briefly with the concept of symbol. This is meant only as a means to clarify of the hermeneutic method.

A symbol discloses the meaning of interpreted existence in that it 'renders manifest the double meaning of worldly or psychical reality' (FP 49/57). Interpretation and symbol are correlated concepts, since 'there is interpretation wherever there is multiple meaning, and it is in interpretation that the plurality of meanings is made manifest' (EH 12/16-17). Hermeneutics approaches the subject as an interpreted existence, situated and acting in a plurality of meaning. In this way, it attempts to clarify the non-coincidence, the conflict, inherent in subjectivity by analyzing the interpretations that the subject performs in order to reappropriate its identity as selfhood (FP 48/56). The nature of symbols explains the structure of interpretation, since symbols render the ambiguity of subjectivity manifest through the multivocity of signs (PD 71/74).

The subject is embedded in a universe of signs. It relates itself to and differentiates itself from the world through language, because it interprets the workings of the world by designating the objects and events in the world. The subject understands that it is situated in a world of meaning by means of signs. I categorize and retain my sensations, plan my actions, and communicate with other persons with the help of signs. The signs most pertinent to human existence are structured in the system of symbols that we know as language. Language is an extremely sophisticated system of symbols that discloses, produces and communicates meaning. A symbol is characterized by a dialectic of presence and absence, since it is bound to sensible reality and, at the same time, transcends it (FP 384/374). Symbols are complex in nature: a symbol, on the one hand, designates an object as this particular sensible object, for example, a stain (to use an example dear to Ricoeur) is a congregation of physical entities, but at the same time it might designate something other than the sensible object, which in this case could be a existential meaning, that is, a stain on one's character (HSP2 312-13/311). The use of symbols in human language covers a vast semantic field that discloses the meaning of human existence. It is through the symbolic nature of language that the intentional structure of subjective experience finds its most clear manifestation of ambiguity: 'the symbolic is the universal mediation of the mind between ourselves and the real; the symbolic, above all, indicates the nonimmediacy of our apprehension of reality' (FP 10/20). Symbols reveal our existential situation as subjects embedded in a meaningful world. They bridge the lacuna between the schematic notion of subjectivity and the 'pathetic dimension' of actual existence. They articulate, through the system of language, the space of multiple meanings in which subjectivity is rooted. The subject finds itself situated in a context of multiple meanings in which it must interpret

its way to understanding and action. Symbols mediate between selfhood and alterity, between the subject and the world, the subject and the other, because they disclose the world as a space of multiple meanings, '[f]or language is the great institution, the institution of institutions, that has preceded each and every one us. And by language we must understand not just the system of *langue* in each natural language, but the things already said, understood, and received.' (TR3 221/400).

The idea that my experience of the world is an experience of a configuration of meanings, i.e. that objects and occurrences relate themselves to me as multiple meanings to be interpreted, yields an important theoretical instrument to clarify and develop the notion of the practical space of reason.

I characterized the practical space of reason as the experience of the world configured as a multitude of values generated by the self, the other, and the world. These values are rooted in the two heterogeneous aspects of the subject, that is, vital (body) and spiritual (reason) desires. However, in the structural analyses the concept of values remained rather unarticulated. To articulate and develop the concept of values, Ricoeur noticed that we have to investigate the feelings involved in our conduct. Our desires transform themselves into values through 'the objects of higher order in which human relations crystallize.' (FM 119/135). Values originate in my desire for well-being (bien-être), basically avoiding pain and augmenting pleasure, but, as we have seen, well-being is a complex affaire when it comes to human subjects. It cannot be reduced to my solitary struggle for survival (FP 471/455), because, due to the heterogeneous aspects in my nature, the objects of my desire transcend the mere physical ones: 'We must add the economic, political, and cultural dimensions to objectivity; they make a human world out of the mere nature they start with' (FM 112/128).

The hermeneutic method explores these objects that constitute the human world through 'the intentional unity of symbols' (FP 522/501). The schematized feelings are constituted around 'the trilogy of the passions of having, power, and worth [avoir, pouvoir, valoir]' (FP 507/487/88. Translation modified; cf. FM 113-25/129-141). That is to say, I assert myself in the world through these three fundamental feelings. My interaction with the world and the other subjects is motivated by these feelings, and values are generated by the interplay of my self-assertion, the presence of the world and the coexistence with the other subjects. So, how can the 'intentional unity of symbols' help clarify the generation of the values that form the practical space of reason?

Ricoeur's thesis is that human action and suffering is primordially mediated through symbols (TR3 221/400). The ambiguity of symbols reveals the complexity of our interaction with the world, since the manifold of meaning disclosed by symbols confirms the heterogeneity of human values: 'The

ambiguity of "things" becomes the model of all ambiguity of subjectivity in general and of all forms of intentionality.' (FP 385/376). The subject does not experience the world merely as a place of an evolving struggle for survival or a continuous maximizing of its own well-being; on the contrary, the world appears as a manifold of meaning. The subject's experience of the world is circumscribed by a complex semantics that cannot be reduced to one aspect of subjectivity, i.e., neither sensibility nor reason. To put it very simply, when the subject experiences action and suffering in its existence, this experience is, for a large part, mediated through symbols.

The human world is a practical space of reason, since it is a space loaded with meaning that concerns the existence of the subject. Ricoeur can therefore claim 'that only man has a world and not just a situation' (MT 201/211). The subject understands objects and occurrences in the practical space of reason not as a succession of separate and non-correlated situations, but as meaningful relations that make up its existence in the world. And this meaning is often confused and ambivalent because of the complex structure of its nature. Human experience is not merely a registration of sensible stimuli, but an interpretation of these. Much of this interpretation is seemly immediate, i.e., instantly, as in the case of perception, but some requires a conscious effort on part of the subject. The instant form of interpretation is a result of the intentional nature of human consciousness, and it can therefore not really be characterized as interpretation (FP 12/21); for example, we instantly and without effort grasp the meaning of a perception. However, the other, more reflective, form is different because it is a 'second-degree intentional structure, which presupposes that a first meaning is set up which intends something, but this object in turn refers to something else which is intended through the first object' (idem. Translation modified).

Now, Ricoeur uses this second-degree intentionality to account for the subject's self-assertion in the world by means of the three fundamental feelings of having, power, and worth. These feelings are the threefold expression of the originating affirmation of subjectivity, i.e., the subject's way of asserting itself in the world. These feelings manifest themselves in the world as objects of meaning that transcend the immediate meaning of perception: 'Such moments are indeed moments of objectivity: to understand theses affective factors, which we name possession, domination, and esteem, is to show that these feelings internalize a series of object-relations that pertain not to a phenomenology of perception, but to an economics, a politics, a theory of culture' (FP 508/488. Translation modified). These 'new' object-relations establish the spheres of meaning that constitute the world of 'the human praxis and the human self' (FP 507/488).

The job of a hermeneutical phenomenology is then to approach the symbolic unity that discloses this human world. Ricoeur proposes a hermeneutic method that combines two seemingly antithetical approaches, the Hegelian and the Freudian. He establishes a dialectic between archeology and teleology in order to approach two fundamental dispossessions of consciousness (FP 460/444-5). Freud and Hegel represent two complementary models of approaching the non-coincidence of the subject. Both of these thinkers conceive subjectivity as different from the immediate self-assertion of the conscious subject. Whereas Freud takes on an archeology of the unconscious to understand the behavior and conscious state of subject, Hegel embarks on the struggle for recognition, the inscription of one consciousness in the consciousness of another, as the aim of the upgrowth of the subject. In this way, we have two apparently opposite directions of analysis. The Freudian that turns towards the internal (unconscious) forces in the subject itself and the Hegelian that aims at the external coexistence of subjects in an intersubjective world (FP 474/458). These opposite directions of analysis, the regressive (archeology) and the progressive (teleology), form 'the two end limits of a single scale of symbolization' (FP 522/501). The unity of symbolization is what we experience as the human world: A space loaded with multiple meanings in which we find ourselves embedded as acting and suffering subjects. The subject is an active part of this unity of symbolization, since it affirms itself as a meaningful existence in the world. Therefore, the hermeneutic method engages in 'the task of appropriating to itself the originating affirmation through the signs of its activity in the world or in history' (NAS 215/219. Translation slightly modified). The subject affirms itself in the workings of the world, and, at the same time, finds itself situated as a part of that world. Subjectivity is movement, action and suffering, and not immediate self-certainty. Accordingly, Ricoeur defines his method as 'a reflective method that has its starting point in the objective movement of the figures of man,' and then further explains that '[r]eflection is the means for deriving from this movement the subjectivity that constitutes itself at the same time that the objectivity engenders itself.' (FP 510/491).

The two end limits of the intentional unity of the symbolization of human actions are, on the one hand, the unconscious, libidinal action of the subject rooted its vital nature and limited by its bodily perspective, its character, and habits; on the other, we have the conscious action that aims at a coexistence with other subjects.

The human world is saturated with symbols that unfold and amplify the concept of value, since they develop and refine the heterogeneous forms of human action and suffering. They secure a conception of human values that is not reduced to either raw vital pleasure (body) or pure spiritual

happiness (reason), but values that are generated as a mixture of both. The hermeneutical method recapitulates and amplifies the structural analyses of human values, since it interprets these values in their manifestation of meaningful objects in the world: 'In hermeneutics symbols have their own semantics, they stimulate an intellectual activity of deciphering, of finding a hidden meaning. Far from falling outside the bounds of language, they raise feeling to meaningful articulation' (FP 19/29). This meaningful articulation is unfolded with the intentional unity of symbolization in that symbols originate by means of the mediation of body and reason in human interaction with the world. In this way, hermeneutics combine the regressive and the progressive direction in the same analysis. In interpreting the symbolization of action and suffering, hermeneutics discloses the meaning of values in relation to human existence as both body and reason. It investigates into 'the figures' of human interaction in order to sort out the complexity of values. The symbols of human values clarify the nature of the identity of subject in the world, because they express, in the objects or the figures of human existence, the lived unity of the non-coincidence of subjectivity.

The hermeneutical approach develops two fundamental aspects of subjectivity that remained rather inarticulate in the reflective analysis, namely temporality and the relation to the other subjects. The dialectic between presence and absence in the symbolic configuration of meaning renders manifest the temporal aspect of subjectivity. The subject is rooted in the present in the sense that it interacts with the immediate context of its being-in-the-world, but, at the same time, it is determined by the non-presence of past and future. The values that motivate its action is not confined to an immediate situation, but part of a world which meaning is derived by what is not immediately given, i.e., actions and sufferings in past and future. The objects with which the subject interacts are understood by means of this dialectic between presence and absence inherent in signs. The symbolic nature of the signs involved in human existence affects the values that shape and determine subjectivity, because the values involved in the subject's relation to the other are refined and developed by the ambiguity of symbols. The subject is not only related to the other in the immediate encounter in the present, but also related over time. The absence of the other does not cancel the value of the other; on the contrary, the value of the other is intact through the signs left by his or her presence. The hermeneutical approach explores this presence of the other in his or her absence through the meaning of the other expressed in the signs that make the world of the subject a human world, namely through words said and deeds done by others. The signs of the other make it explicit that the values involved in my existence depend on more than my own well-being: 'My existence for myself is thus dependent on the way I am regarded by other people; the self is shaped

by the opinion and acceptance of others […] It is through the medium of these works or monuments that a certain dignity of man is formed, which is the instrument and trace of a process of reduplicated consciousness, of recognition of the self in another self' (FP 523/502-3).

Ricoeur's turn to hermeneutics is an attempt to confer the structural analyses of the stripped notion of subjectivity the pathetic dimension of lived existence. The subject is thus approached as a living subject in time, intertwined with other living subjects. The foregoing analyses have outlined the notion of a practical space of reason wherein the subject experiences itself embedded in a space of values issued from the world, the other, and itself. The notion of affectivity showed that the subject in its inmost core, the heart, is marked by a non-coincidence, which prevents immediate identity of the subject. Subjectivity is characterized by a constant conflict between two fundamental aspects of the self, sensibility and reason, experienced as vital and spiritual desires. The subject does not possess itself, but must become itself in the mediation of the conflict between selfhood and alterity (FP 45/53). This conflict cannot be resolved at the level of structural or reflective analyses, since the subject has to find itself through the alterity that makes up its being. In order to investigate into this alterity, the analyses must turn to the objects and persons that qualify the being of the subject as an coexistence with other subjects situated in a human world. The identity of the subject must be retrieved by means of an interpretation of its existence, i.e., the story of the subject's affirmation of itself in the world through interaction with the world and the other. We have seen how this interpretation is bound to the ambiguity of symbols, which, on their part, reveals the complexity of the values that constitute the practical space of reason. Values have multiple meanings. Nevertheless, the meanings involved in the symbolization of the subject's interaction with the world are not arbitrary, but gathered in an intentional structure that is intrinsically related to the nature of subjectivity. And particularly two aspects of the semantics of symbols are of interest to a theory of subjectivity, namely temporality and the relation to other subjects. The meaning of human values is not bound to immediacy, the present, but extends itself back into the past and forward into the future. Furthermore, values cannot be reduced to the unconscious, libidinal workings of one individual subject (archeology), but must include the conscious inscription of the actions of one subject into the actions of another (teleology).

To clarify the problematic identity of the subject in relation to these two aspects of the symbolic interaction with the world, Ricoeur introduces another concept in his analyses, namely the concept of narrative identity. Here I will merely introduce the concept, that is, briefly explain the relation

between hermeneutics and narrative, and thereby prepare the use and development of the concept in the remaining sections.

Explaining the hermeneutical turn, Ricoeur says that '[t]he man who speaks in symbols is first of all a narrator; he transmits an abundance of meaning over which he has little command […] interpretation consists less in suppressing ambiguity than in understanding it and explicating its richness' (FP 49/56).

Narrative Identity

The concept of narrative identity is first introduced in relation to the temporal character of human experience, more specifically in 'the mediations between the discordant concordance of phenomenological time and the simple succession of physical time' (TN3 22/42; cf. 90-91/167-169). Human existence is marked by time, and, in order to develop the schematic notion of subjectivity into concrete notion of lived subjectivity, one must clarify how the subject is characterized by time[12]. Ricoeur challenges what he calls the aporetics of temporality. Human temporality cannot be reduced to neither personal, phenomenological time nor impersonal, cosmic time (TN3 20/40-2). Once again he confronts himself with (among others) Kant and Husserl. Whereas Kant emphasizes invisible time as an objective transcendental condition for conscious experience, i.e., time of nature, Husserl investigates the phenomenological time of intentionality, i.e., subjective time experience. Ricoeur, however, retains that neither can fully explain the temporal character of human experience because their respective analyses exclude that of the other (TN3 57/106-7). On the contrary, the particular nature of human time depends on both the subjective and objective conception of time and finds its explanation in what Ricoeur calls 'a third-time' (TN3 245/441; cf. 96/177). This third-time is time configured both by what affects the subject involuntarily and by how the subject acts voluntarily, and these two aspects come together in the practical category of 'narrative identity' (TN3 246/442). The two aspects of narrative identity are interlaced in a dialectic between historical and fictional narratives: 'From these intimate exchanges between the historicization of the fictional narrative and the fictionalization of the historical narrative is born what we call human time, which is nothing other than narrated time' (TN3 102/185). History is bound to what has really happened, a concrete event that occurred in a specific physical place and time. History is something that cannot be undone, reversed or reflected away.

[12] 'In this regard my basic hypothesis is the following: The common character of human experience, which is marked, articulated, and clarified by the act of storytelling in all its forms, is its *temporal character*' (NPH 63).

And yet history, since it is based on an interpretation of archives, documents, and traces, always contains an imaginary element, or what Ricoeur names a 'fiction-effect' (TN3 186/337) or 'quasi-fictive' element (TN3 190/346). History is a narrative told as a 'task of memory' (TN3 189/342) in order not to forget the actions and sufferings already done. We unfold this narrative by engaging in imaginative mediations on cosmological time (exemplified by the calendar and the sundial) and phenomenological time: 'History, I said, inscribes the time of narrative time within the time of the universe' (TN3 181/331). We tell history to understand human time, and, not less important, we must understand time in order to understand the nature of humanity: 'To think of history as one is to posit the equivalence between three ideas: one time, one humanity, and one history' (TN3 258/461). Fiction, on the contrary, is the reenactment of what might have been, of 'the probability of the universal' (TN3 191/345). When we engage in fictive narratives we explore 'certain possibilities that were not actualized in the historical past,' and therefore fiction is 'able, after the fact, to perform its liberating function' (TN3 191/347). Fictional narratives engage in a creative imitation, a mimesis, that liberates the logical structure or meaning of the possible events in a human life; or said otherwise, 'I offer to myself to the possible modes of being-in-the-world which the text opens up and discloses to me' (MPH 177). Fiction possesses what Ricoeur calls a 'heuristic force' to re-describe reality as we know it (IDA 248).

History and fiction are interwtwined in a circular relationship in the act of narration (TN 3 190/345), which refigures the experience of time in such a way that we come to terms with the aporetics of temporality. By narrating time, lived time (phenomenological time) is reinscribed in the time of nature (cosmic time) (TN3 99/181-2) so that a bridge is set over the polarity of the two conceptions of time (TN3 244/439). Human time is neither lived time nor the time of nature, but a complex of the two. We are bound to the time of nature through our body, since our body is firmly rooted in the physical workings of the world, and still, lived time is not restricted to the physical aspect of time. Our past is not only measured by the anonymous quantity of days gone by or by the traces of time visible on our faces (TN3 138/245-6); it also reveals itself as the intrinsic time of consciousness that we experience in memory. We are not situated in time as helpless spectators; on the contrary, we are involved with time as agents in the constitution of time. Human time springs form the ambivalence of acting and suffering (TN3 221/400) in the sense that we, on the one hand, are affected by time through the evolving of anonymous cosmic time, the succession of generations, cultural tradition, and the involuntary events of a life-story. On the other hand, we ourselves constitute time in that we seek to understand and ordinate the experience of temporality. We

develop such an understanding of time in different ways, for example, by measuring time with the clock and the calendar or telling it through history and fiction. Time becomes intelligible through our attempt to speak about our experience of temporality, or said in another manner, we understand our relation to time and how we, as human beings, are situated in time in speaking about time by means of narratives and metaphors: 'Humanity becomes its own subject in talking about itself' (TN3 212/383). Thus the third-time, human time, is time as it is narrated with reference to the identity of the subject who speaks about and in time.

Temporality is a critical for the question about personal identity, or more precisely, we cannot inquire into the identity of the subject without dealing with the temporal dimension of human existence (OSA 113-4/137-8). We are affected by time and try to come to terms with time because time profoundly affects our identity as humans. Time is the horizon within which we live, act, and suffers, since (and here Ricoeur joins and, at the same time, criticizes Heidegger's analyses of time) human existence is characterized by being 'intra-temporal', by its historicity (TN3 122-3/222-3)[13]. Nowhere does the non-coincidence of the human subject become so critical as in the temporal character of a human life. A person changes over time, gains new character traits, transforms and loses old ones, and yet she remains the same person. Still, it is problematic simply to assert that the person remains the same person over time, since, as we have seen, personhood is not something static or certain, but a conflict approached in terms of an intrinsic process of appropriation and affirmation (or restoration) of the self.

The concept of narrative identity is a part of the hermeneutical approach to this conflict inherent in human subjectivity. The hermeneutic answer to the question 'who am I?' understands human existence as an encounter between the subject and the world. The world becomes a human world through the mediation of signs in that 'human acting is intimately articulated by signs, norms, rules, and evaluations that situate it in the region of meaning, or, if you will, within the symbolic dimension' (TN3 232/419). The subject engages with the world, with its own identity in the workings of the world, by means of symbolic mediations. The meaning of the world, and thereby the meaning of the individual subject in the world, is apprehended by an interpretation of the symbolic unity that discloses the world as a human world, a practical space of reason, wherein the subject continuously seeks to come to an understanding of what affects it (alterity) and what it does itself (selfhood).

[13] According to Ricoeur, Heidegger pays little attention to "ordinary time [le temps vulgaire] and within-time-ness", which Ricoeur sess as '*time of the trace* [*le temps de la trace*] that we find in, for example, calendar time (122-3/22-3).

The concept of narrative identity is introduced to cope with the problematic temporal character of subjective identity in that the concept integrates the apparent polarity in the identity of the subject, between permanence and diversity over time (OSA 140/167-8). How can I affirm my self-constancy, or deny it for that sake, if I do not make my 'life itself a cloth woven of stories told' (TN3 246/443)? I relate myself to the identity of my existence in time by examining my life as a complex of stories told about my life. Thus, Ricoeur establishes a connection between self-constancy and narrative identity that, as he claims, confirms one of his oldest convictions: 'The self of self-knowledge is the fruit of an examined life […] And an examined life is, in large part, one purged, one clarified by the cathartic effect of the narratives, be they historical or fictional, conveyed by our culture. So self-constancy refers to a self instructed by the works of a culture that it has applied to itself' (TN3 247/443-4). The identity of the subject cannot be reduced to what happens to the subject over time, to the way time works on the subject (the cosmological time), but must necessarily include how the subject itself acts on its being affected by time (lived time). The dilemma of permanence and diversity, i.e., whether the subject remains identical to itself through the diversity of different temporal states or the identical subject disintegrates into a mere flow of different 'nows' with nothing but a contingent connection, disappears when we approach the identity of subject as 'refigured by the reflective application of narrative configurations' (TN3 246/443. Translation slightly modified).

As noted above, narrative identity is a practical category, which means that it deals with aporetics of time and identity in relation to the concrete existence of the subject, that is, with how the subject acts and suffers in the actual coexistence with other subjects in a human world. Ricoeur would in no way deny that narrative identity has its limits with regard to the explanation of both time and identity. In fact, he explicitly points to the limits of narrativity in relation to the unrepresentability or inscrutability of time (TN3 270-3/482-8), and further emphasizes that '[n]arrative identity thus becomes the name of a problem at least as much as it is that of a solution […] and has to link up with the nonnarrative components in the formation of an acting subject' (TN3 249/446-8). Hence, Ricoeur insists that the concept of narrative identity has to be understood in relation to the other aspects of his theory of subjectivity. In other words, it is a way to deal with the problem of identity as it is felt and lived by individual subjects in a human world, or, to put it otherwise and more in line with our previous analyses, it is a theoretical instrument coined to develop the basic conflict between self (vital desires, i.e., body) and self (spiritual desires, i.e., reason) on the level of concrete subjectivity. It is a part of his hermeneutical turn in the sense that with this concept Ricoeur seeks to

reinstate the schematic notion of subjectivity in a concrete, historical context. We cannot understand human subjectivity in isolation from its existence in time, and time cannot be reduced to neither anonymous cosmic time nor lived phenomenological time; it has to be a complex of both as we actually experience it, living in a human world made up, for a large part, by history and fiction. The concept of narrative identity conceives the human subject as a being situated and embedded in a world of values that are generated by the subject itself and by that which is not the subject. The identity of the subject is a practical matter. We cannot understand it without taking into account the historicity of the world in which the subject is embedded. On the one hand, this historicity 'is itself the record of human action' (MT 207/218) as we experience it through traces such as archives, documents, generations and traditions. We create a past by narrating the times past, which, however, does not mean that we are free to create an arbitrary past based on our likes or dislikes, since '*we belong to history before telling stories or writing history*' (NF 294). On the other hand, the historicity of our being-in-the-world is not restricted to an understanding of the past, but expanded and thickened by fictional narratives that do not limit themselves to re-describe what has been. On the contrary, they liberate our imaginative capacities in that they imitate the actual possibilities hidden in the ontological structure of human action and thereby point towards the future, or said differently, 'what mimesis imitates is not the effectivity of events but their logical structure, their meaning […] Mimesis is a kind of metaphor of reality' (NF 292).

Interpreted in this way, we see how Ricoeur uses the concept of narrative identity as a way to articulate and systematize the insights developed in his hermeneutical turn, which, on their part, draw heavily on the structures uncovered by the reflective-phenomenological analyses of the schematic notion of subjectivity. The hermeneutic approach to subjectivity elaborates and thickens the practical space of reason, since it emphasizes the historicity of values and the unyielding presence of the other subjects.

In the remaining sections, I will develop the concept of narrative identity in order to clarify how Ricoeur uses the concept to confer historicity to the schematic notion of subjectivity, or said differently, how he reinstates the subject in the pathetic dimension of situated existence. Furthermore, I shall argue for the necessary connection between the hermeneutic approach to subjectivity and ethics, since the conflict between selfhood and alterity becomes more critical when the existence of the subject is considered as concrete existence and its values as embedded and situated in socio-cultural context. As mentioned above, I inverse the structure of the analyses done in chapter one so that I start with affectivity, then turn to action and end with ethical experience.

The three sections are structured in this way to emphasize how narrative identity is rooted in the 'nonnarrative components' of human subjectivity. I place the analysis of narrative identity in between the two nonnarrative dimensions of subjectivity, affectivity and ethical experience. I begin with affectivity as it is experienced in bodily experience. The narrative nature of human identity is bound to the embodied nature of subjectivity, which means that the subject already *exists* in the world in a certain way *before* it begins to understand this existence through narrative configurations. In the following section about human action, I will deal explicitly with the concept of narrative identity in order to clarify, on the one hand, how Ricoeur uses this concept to examine the question of human identity in the concrete world of history and culture, i.e., his attempt to redress the stripped notion of subjectivity with the properties that he intentionally abstracted from in the structural analyses, and, on the other hand, to evidence how human identity necessarily transcends the capacity of narration and leads to the nonnarrative dimension of ethics. The final section will, in way of conclusion, focus on the relation between subjectivity and ethics. I will try to substantiate his claim that the question of subjectivity remains obscure and inconsistent if it is not dealt with in relation to the notions of humanity and personhood.

The Affected Subject: Body and Alterity

One of the most pertinent ideas in Ricoeur's works is that the body is the primary existential mediation between self and world (OSA 150/178). Without the body there would be no interaction between self and world, since the subject could not inscribe itself in the world nor be affected by the world (OSA 319-20/370-1); world and self would remain two separate beings. Hence, the body remains fundamental at all levels of the analyses of human subjectivity. But the body is more a problem than a fact. The body is constituted as 'the mixed structure of "I – such a one"; as one body among others, it constitutes a fragment of the experience of the world; as mine, it shares the status of the "I" understood as the limiting reference point of the world. In other words, the body is at once a fact belonging to the world and the organ of a subject that does not belong to the objects of which it speaks' (OSA 54-5/71-2. Translation modified; cf. FN 19, 193). Ricoeur uses the analyses of the body done in the phenomenological tradition (Heidegger in particular) and confronts the problem that one's own body is at once 'the flesh's intimacy to the self and its opening to the world' (OSA 326/377)[14]. In the subjective experience of one's own body, Ricoeur finds the origin of all experience of alterity: we are exposed to that which is not ourselves through our body.

[14] Ricoeur adopts the Husserlian distinction of the experience the body as *body* (*Körper*, a physical object among other physical object) and *flesh* (*Leib*, my body).

Therefore, it is through our bodily experience that we first become aware of the passivity inherent in being a person. Alterity is that which transcends or even eliminates our capacity to act, to set our mark on the events of the world, and thereby it reveals a suffering passivity in the existence of the subject (OSA 319-20/370). The concepts of non-immediacy and passivity are disclosed by the alterity involved in the embodied nature of subjectivity; they find their first (personal) expression in the ambivalent experience of one's own body.

Now, Ricoeur wants to develop the problem of the alterity in subjectivity, the conflict between the self as reason and the self as body, at a practical level. In fact, he is convinced that the question of human subjectivity cannot be resolved but by turning to the practical dimension of selfhood (OSA 111-2/135-6)[15].

Non-immediacy and Ambivalence of the Body

We have seen that it is by means of the body that the subject is capable of affirming itself in the world and, at the same time, that the body restricts the subject in this affirmation. But whereas the previous analyses limited themselves to the structural analyses of the subject's intentions and projects, Ricoeur now try to develop an ontology that places the subject in the world of history and culture and, furthermore, explain the relation between action and subjectivity (OSA 74/93); or said otherwise, an ontology of 'a being in project' (OSA 86/107. Translation modified). This passage from phenomenology to ontology must first deal with the enigmatic nature of the body in subjectivity, i.e., one's own body (OSA 319/368).

An ontology of selfhood is rooted in the ambivalence of the body, because, on the one hand, my body is mine, i.e., a part of my most intimate experiences and the exteriorization of my intimacy in being the medium of my inscription in the surrounding world; on the other, my body is a part of that which is not mine, i.e., the alterity experienced in being subjected to physical causality and the impersonal events operating in the world (OSA 322/372). Ricoeur formulates this fundamental ambivalence in the following way: 'To the extent that the body as my own body constitutes one of the components of mineness, the most radical confrontation must place face-to-face two perspectives on the body – the body as mine, and the body as one body among others.' (OSA 132/159).

[15] As he writes elsewhere, claiming his affinity with Aristotle: 'In this way the practical judgement gives a rich content the idea of the human *ergon*' (CR 88/134). And this human *ergon* is the peculiar way in which human beings interact with the world in transforming their mere presence into existence, into a life: 'Taken as a singular term, the word "life" receives the appreciative, evaluative dimension of *ergon* which is used to qualify man as such. This *ergon* is to life, taken in its entirety, as the standard of excellence is to a particular practice" (OSA 177-8/209)

The non-coincidence in the core of subjectivity finds its primary experience in the ambivalence of the body. Our decisions must be worked out by means of the body to become actions and, further, the world reveals itself only through our bodily senses. I cannot escape my body neither by abstracting from it nor by ignoring it. That is, my decisions might be done without paying attention to my bodily constitution, but my attempt to realize such decisions my own body reveals the limits of my possibilities by transforming them into *actual* possibilities. My practical knowledge of the world is shaped by the bonds of my senses and the abilities of my body, and my self-knowledge must be balanced with my practical knowledge of the world. That is to say, I become aware of my being this individual subject with these specific capacities and situated in this particular world through the appropriation of my own body. In bodily experiences, the subject experiences an alterity that is part of itself, and thus the subject cannot dispose of it, but is, by plain fact, constricted to integrate this experience in the attempt to understand itself. For example, I cannot see what my eyes look like when I am nervous.

We may be able, on a theoretical level, to abstract from our bodily constitution, but all practical experience is bodily experience, since the subject acts and is affected through its bodily constitution. I enter the world through my body, since my actions, which are carried out by means of my body, determine my being-in-the world. Thus, my body assumes an instrumental character by means of which I liberate myself from the pure passivity of not-acting. I set my mark on the events in the world (OSA 319/370), which means that the body that I find my self bound to, my flesh, is *my* body in the most intimate way. Still, the body also escapes my initiative in that it situates me in a given context, disobeys me, confines me within certain limits and determines the scope of my actions. In this sense, the world works on me, affects my being this individual subject through my own body. I am bound to a certain perspective and given this particular physical constitution (e.g., these ears, this hair, this color of skin, and this tone of voice) without my own consent; I am subject to pain, sickness, ageing, and eventually death because of my body. Or said even more simply, my body does not always behave the way I want it to behave. For example, I reach out for a cup of coffee and end up with a pen in my hand. Thus, I live 'with the otherness of the flesh that I am' (OSA 326/377)

However, the ambivalence of the body is not limited to the neat distinction between praxis and pathos, acting and suffering, but includes the dialectics of selfhood and alterity in action itself, i.e., in the way I use my body to realize the affirmation of my selfhood. The moment I transform my decision in action, that is, inscribe my intimacy into the world, I am no longer in complete

command of my decision. I cannot control how my actions inscribe themselves in the events of the world. My body opens up to the world, but does not have power over the world, since it is an integrate part of the workings of the world; the scope and effect of my action necessarily escapes my initial intention (OSA 294-5/341-2).

In my bodily existence, the non-immediacy of my being is experienced as a facticity, i.e., a practical condition of my selfhood. The person that I am is not something immediately given. I must gain control over my body, pay attention to how my body inscribes my being in the world and, moreover, I have to accept that my body is not only an object of my will, but also a part of that over which I have no control. My actions are realized through *my* body, but already part of an alterity in the action itself (because of the alterity in my body). To become the self that I am I must accept and consent to the being of my body. My body is what opens up to the world, my particular access to the world, and, in this way, it transforms the world from something totally foreign to the primary alterity of human finitude 'insofar as it is sealed by embodiment' (OSA 327/378). Or put differently: 'one's own body is revealed to be the mediator between the intimacy of the self and the externality of the world' (OSA 322/372). The embodied nature of human subjectivity shapes the experience of the world and the self-experience of the subject in such a way that the world in which the subject is situated is transformed from something totally foreign, something without relation to the being of the subject, to an alterity that is part of the subject in its existence in the world. The world becomes something that the subject cares about: 'Only a being that is a self is *in* the world; correlatively, the world in which this being is, is not the sum of beings composing the universe of subsisting things or things already-to-hand. The being of the self presupposes the totality of a world that is the horizon of its thinking, acting, feeling – in short, of its *care*' (OSA 310-1/360).

Therefore, to clarify how the world and the subject together configure the world of the subject (the practical space of reason), we have to consider, as an inevitable first step, the passivity inherent in the bodily experience of the world and, more importantly, how this passivity determines how the subject experience its being-in-the-world.

Passivity and an Ontology of the Flesh

In the last section of chapter one, we saw how the affective dimension of subjectivity revealed, through the complex nature of human feelings, the non-coincidence in the heart of the subject. Through our feelings, we continuously, at an intimate and pre-reflective level, experience a conflict between the original affirmation and the existential negation (perspective, character, vital desire)

inherent in human subjectivity. Now, Ricoeur wants to develop the implications of this fundamental conflict at a practical level. That is to say, when we are affected *through* and *by* the body, how does the body determine our practical experience?

First, the fact that we are affected *through* the body, i.e., the body is our openness to the world, reveals that we are the sole origin of our existence. We do not experience our coming into the world. Our existence is marked by an experience of 'already having been born and of finding oneself already there' (OSA 327/378). In this way, we find ourselves unwillingly present in and exposed to the world through our embodiment. Ricoeur elaborates on the Heideggerian concepts of *thrownness* (Geworfenheit) and *mood* (Befindlichkeit) in that he explores what it means to the subject to interact with a world in which it has not put itself. The affective dimension of this being-already-there is experienced by the subject as a certain affective state, a mood (Befindlichkeit), i.e., a feeling of being this individual subject in this specific world. Heidegger's analyses of thrownness and mood opens up for an ontology of the flesh, that Heidegger himself, in Ricoeur's opinion, never developed: 'an ontology of the flesh, in which the latter gives itself to be thought not only as the embodiment of "I am" but as the practical mediation of that being-in-the-world that we are in each case. This conjunction between flesh and world is held to allow us to think the properly passive modalities of our desires and moods as the sign, the symptom, the indication of the contingent character of our insertion in the world' (OSA 326-7/377-8, note). Further, in an extreme sense, the contingency of our existence in the world is a burden that the subject has to endure in 'the task of having-to-be' (idem), since it has not itself chosen to exist. Therefore, our body, which is itself rooted in passivity, places us in the world in a certain way. Our bodily existence precedes every distinction between voluntary and involuntary because that 'I will' is firmly rooted in 'I can'. I only have the possibility to choose, to affirm myself in the world, because I am *given* certain bodily capacities in the first place (OSA 324/375).

Now, how does Ricoeur develop this ontology of the flesh that he finds so important for the notion of subjectivity? His late analyses in OSA leaves the reader rather perplexed, because he repeatedly returns to the ontology of the flesh and emphasizes the importance of what he calls authentic spatiality as a crucial element of the coining of such an ontology. And still, he never arrives at an explicit definition of what he intends by an ontology of the flesh. He leaves the reader only with sporadically suggestive hints to the notion of authentic spatiality. For example, authentic spatiality is essentially different from 'the geometric space as a system of indifferent places' in that it is 'the spatial dimension of being-in-the-world […] the backdrop of the spatiality of available and

manipulable things' (OSA 328/379). Or, he writes that the self is in relation to the world as to a totality of concerns in the sense that everything concerns the subject (OSA 314/363). And further, that we need to coordinate the human initiative with the 'the movement of the world and all the physical aspects of action,' because only by doing so can we explain the extension of the totality of our concerns from being-alive to the praxis of living well (idem). Still, he does not gather these observations into a systematic notion, and, therefore, the importance of the reflections on authentic spatiality and an ontology of the flesh remains somewhat inarticulate.

Here, though, the previous analyses of the schematic notion of subjectivity turn out to be helpful. The investigations done in the early works, FN and FM in particular, have already delineated such an authentic spatiality in relation to the practical space of reason.

The practical space of reason is, as stated (perhaps too) many times now, the conscious space of experience produced by the transcendental imagination in the mediation of reason and sensibility in such a way that the different motives and their joined feelings are weighted against one another according to a hierarchy of heterogeneous values generated by the world, the self and the other. At this point, however, Ricoeur wants to develop this schematic notion of subjectivity by grounding it in the historicity of concrete existence by means of a hermeneutic-phenomenological method. He wants, so to say, to redress the stripped, structural notion of subjectivity with the concrete, pathetic dimension of the subject's actual (co)existence with other subjects. This shift in interest (and therefore, in consequence, also in method) might be the reason why Ricoeur does not develop the ontology of the flesh that he nevertheless believes so important for an ontology of selfhood (OSA 328-9/379-80). It might be due to his emphasis on a more hermeneutic (i.e., concrete, historic, and practical) notion of subjectivity that the immediate bodily experience remains inarticulate in the later works. I believe, however, that the idea of an ontology of the flesh is important for Ricoeur's general theory of subjectivity and that it could become more articulated and clear, if one puts the somewhat scanty analyses of bodily experience in OSA in relation to previous analyses of the schematic notion of subjectivity.

We find ourselves embedded in a world that we have not chosen to be a part of. We, as individual subjects, are different from the world, and still we are a part of the world in which we conduct our existence. Our existence in the world is characterized by the continuous experience of both selfhood and alterity, and neither world nor subject can be fully understood in isolation from the other: 'There is no world without a self who finds itself in it and acts in it; there is no self without a world that is practicable in some fashion' (OSA 311/360). This, however, is not the same as to say that it

is nonsensical or illegitimate to consider the world in abstraction from the subject, as it is done, for example, in physics and other branches of the empirical sciences, but only that the subjective experience of the world, i.e., the world of the conscious subject, cannot be reduced to the impersonal world uncovered by the physics or biology. As we saw chapter one, Ricoeur is interested in the interaction of the world and the subject, and this interaction takes place in what the subject experiences as the practical space of reason. The subject experiences the world personally in the sense that the world is a configuration of heterogeneous values that correspond to the subject's concerns, or expressed differently, the subject cares about the world in which it finds itself embedded.

The previous analyses of values and motivations help clarify the notion of authentic spatiality, since they place bodily experience in the larger context of the practical space of reason. The ambivalence of the body, marked by activity and passivity, selfhood and alterity, is the primordial (i.e., immediate) experience of the interaction of subject and world. We experience the world through our interaction with the world in such a way that the world becomes the correlate to our actions and our sufferings. Thus, authentic spatiality is the immediate, pre-reflective experience of world as a practical space of reason, that is, as a configuration of heterogeneous values that express the totality of our concerns. It is a spatiality internal to the flesh, because the primordial form of sensing places the subject in a pre-linguistic space that becomes a practicable world through the subject's capacities to interact with the alterity present in that space (OSA 325/376). The subject is part of that alterity, because its body is merely a body among others, and yet it is capable of appropriating the alterity of the space in which it finds itself, transform it into a familiar world by means of the capacities to affirm itself inherent in its own body, its flesh. The ontology of the flesh is the first, necessary step toward an ontology of selfhood, since it gives an account of how the being of the subject is related to the being of the world. It is the spatial equivalent to the temporal analyses in TN3 in the sense that the subject's bodily experiences (perception and voluntary movement) is inscribed in the anonymous workings of the world in the same manner that lived, phenomenological time-experience is inscribed in cosmological time (OSA 326/376-7). Thus, an authentic spatiality is the central concept in an ontology that brings to the forefront the complexity of embodiment, that is, the fact that the subject is both part of the physical world (spatially and temporally) and, at the same time, somehow different from the workings of that world. In our immediate interaction with the external world, we feel and experience the dialectic of selfhood and alterity that is internal to our being human subjects (idem). Although our bodily experience might

seem immediate, it is still grounded in a more primordial mediation between the two fundamental aspects of human subjectivity, namely body (passivity) and reason (activity).

A hermeneutical approach to the notion of subjectivity is not possible without the preliminary reflective-phenomenological analyses of the structural properties of subjectivity, since these clarify the transcendental, pre-reflective structures of subjective experience that remain primary to a more concrete, hermeneutic notion of subjectivity. Ricoeur has always emphasized this, but, at the same time, he insists that the phenomenological approach finds its limit, when it comes to reinserting the structural notion of the subject in the world, i.e., developing an ontology of the self (SPH 144; PW 79-80; KH 245-6; FN 3-4/7-8, 471-2/443-4; FM 45-6/63; PH 101/44, 114/61). The structural analyses of the subject ends in the ontology of the flesh, since the experience of my own body points to the limit of my self-awareness, to the passivity in my-being-this-body, 'namely, not what it means that a body is my body, that is, flesh, but that the flesh is also a body among bodies. It is here that phenomenology finds its limit' (OSA 325-6/376-7). Phenomenology cannot account for the alterity experienced in human subjectivity, since this alterity exceeds the conscious capacities of the subject. The hermeneutic approach to subjectivity therefore begins where phenomenology ends, namely with the conflict between alterity and selfhood in the heart of the subject in that it emphasizes the mediated nature of subjectivity by means of interpretation.

Therefore, in my opinion, the structural analyses of the early works are precisely the outline of such an ontology of the flesh, because they clarify our basic bodily interaction with the world. The schematic notion of subjectivity provides an indispensable fundament, by outlining the structural elements of the practical space of reason, for the development of an ontology of selfhood, since it insists on the primordial dialectics of selfhood and alterity inherent in the most basic form of subjectivity, namely that of bodily experience. It is the first step to solving the apparent paradox that 'if there is a being of the self – in other words, if an ontology of selfhood is possible – this is in conjunction with a ground starting from which the self can be said to be *acting*' (OSA 308/357).

Before turning to the problem of selfhood and identity in action, we might, for the sake of clarity, sum up the main points of the analyses done in this chapter so far:

- Ricoeur changes (around 1960) his method from reflective-phenomenological to hermeneutic-phenomenological in order to confer the schematic notion of subjectivity the pathetic dimension of concrete historic existence. One can say that he redresses the stripped subject with the problems that he intentionally abstracted from in the early works (temporality, culture, historicity, and the concrete presence of the other).

• The hermeneutical turn thickens and complicates the notion of the practical space of reason, since it emphasizes the existence of the subject in time and culture. The concepts of motivation, value and the other become somewhat more complicated in that they are faced with the contingency of concrete existence, i.e., the subject as situated in a certain social, linguistic, and cultural context.

• However, although the situated subject is determined by the social and cultural context in which it finds itself, Ricoeur is cautious about not reducing the nature of the subject to these more contingent factors. He therefore emphasizes the embodied nature of subjectivity as a necessary presupposition for the hermeneutical analysis of the situated subject. This presupposition is found in his analysis of the schematic notion of subjectivity. Here he developed an ontology of the flesh that clarifies the pre-reflective structures of subjectivity with regard to which the further analyses of the situated subject must be carried out.

Action: Selfhood and Identity

How, then, does Ricoeur deal with the question of selfhood and identity in situated subjectivity? Already in FN (1950), Ricoeur was convinced that the phenomenological description of the structures of subjectivity must be adjoined with metaphysical or ontological analyses in order to clarify 'the unity of man with himself and his world' (FN 467/439). Although the affective dimension of subjectivity is what 'truly represents man's humanity' (FM 136/152), the conflict solicited by vital and spiritual desires cannot be resolved but through an analysis of the social and cultural objects in which the subject finds itself and affirms itself in the situated nature of its existence. Thus, Ricoeur is concerned with the life of the subject, but the term life 'is not taken in a strictly biologic sense but in the eticocultural sense […] The word "life" designates the person as a whole in opposition to fragmented practices' (OSA 177/208-9).

The structural analyses emphasized that the subject exists as a self only to the extent in which it acts. Action remains a salient feature in Ricoeur's late analyses. This time, however, the question of action is not a strictly theoretical one about the intrinsic structures in human action, but is transferred to a practical level (CR 138/90): How does the self relate to its action? And further, how do we distinguish a personal action from an impersonal event?

In order to answer these questions, we need an ontology that accounts for the difference between action and event. If we adopt a materialist ontology that reduces personal actions a subclass of impersonal events, we cannot distinguish between action and event and thereby not articulate the relation between agent and action (OSA 74/93). By assigning personal action to an ontology of

anonymous events, we neglect the personal aspect that is, according to Ricoeur, the fundamental aspect of human action, namely the question of "who" did the action (OSA 60/78). Therefore, we must 'seek in another sort of ontology, one more consonant with the search for the self, the genuine place of linkage between the action and its agent' (OSA 74/93). Ricoeur is on the look for an ontology that does not sacrifice the specific difference between the nature of human and non-human events in order to gain a clear and unified explanation of everything that happens in the world (OSA 83-4/104). We have to respect the fact that there is something substantially different at stake when a person brings about a change in the world with respect to an impersonal event such as an earthquake or a dog causing an accident. Nonetheless, Ricoeur does not ignore another pertinent fact about personal action, namely that, although a personal action is different from an anonymous event, the effectuated action is still a part of the physical universe and therefore sometimes very difficult to distinguish from a mere event (OSA 69/88). So, this new ontology has to pay attention to both the personal and impersonal aspect of action, but most importantly it must be differentiated from an ontology based strictly on impersonal events: 'This different ontology would be that of a being in the making, possessing de jure the problematic of selfhood, just as the problematic of sameness belongs de jure to the ontology of events' (OSA 86/107; cf. OSA 96/118).

Motives and Causes

What characterizes, intuitively, the difference between personal actions and anonymous events? One might say intuitively, the idea or feeling that an action could have been done otherwise or even prevented by the agent, whereas an event could not have, i.e., it is not a matter of doing or not doing, it simply happens. In our attempt to find cause for an occurrence in the world, our search stops with the agent when it comes to action (he or she did it, and that's it!), whereas we face all sorts of difficulties trying to narrow down on an ultimate cause for anonymous events (this happened because of that which again was caused by this, and that, on its part, was a reaction to this, and so forth *ad infinitum*). Said in another way, personal action is characterized by an intention: 'Describing an action as having been done intentionally is explaining it by the reason the agent had to do what she did' (OSA 75/95).

Now, a personal reason for an action cannot be explained by the causal model built on an ontology of anonymous events, and the key-notion for an alternative explanation is the concept of motivation (OSA 77/97). As we saw in the previous chapter only human beings have motives and not causes (p. 26-30). Ricoeur insists on this human particularity and repeats that '[o]ne has not forgotten the

opposition between the event that happens and the event that one brings about or the opposition between cause and motive' (OSA 103/125. Translation slightly modified).

A motive is what connects the action and the agent. The action is ascribed to the agent in that the agent must appropriate the action. The agent cannot dispose of the action made, but remains the possesor of what he or she makes (CR 97-8/148-9). The action causes a change in the world, which cannot but finds its origin in the self that initiated that change (OSA 109/133). The "how" and the "why" finds its common root in the complex question "who?" (idem). The opposition between cause and motive finally finds its justification in the distinction between subject and object. An action differentiates itself from an event in the way that its cause can never derive from an object, but do always originate in a subject that possesses the capacity to realize or restrain from that given action. The subject may be an object of external forces, and yet, it always remains a subject because of its capacity to deliberate on what to do when it is affected by the world in which it is situated: 'Ascription consists precisely in this reappropriation by the agent of his or her own deliberation: making up one's mind is cutting short the debate by making one of the options contemplated one's own' (OSA 95/117). That is why we are directed to the self and its motives and not to the forces of an anonymous causality, when we want to know the "how" and "why" of a certain action.

The previous section showed that the distinction between selfhood and alterity in subjectivity is not as clear-cut as might seem in the distinction between motive and cause. The self is part of the world and as such it is always difficult to discern exactly to which extent one may say that an action depends on the self. Nevertheless, Ricoeur emphasizes that in order to talk about subjectivity, we must not let the nature of action be confounded with the causality of anonymous event, because '[t]he action must be able to be said to depend on the agent in order to be blamable or praiseworthy' (OSA 101/123).

The nature of personal action is therefore to be sought in the nature of the self that initiates the action. Once again, we have to take a closer look on the self that is capable to cause an event in a world governed by laws different from those of the self. Thus, we must individuate the dimension of personal action within the anonymous events in the world by restricting our analyses to where we can talk about responsibility (OSA 106-7/130). This is not at all an easy process, but we may begin by delineating 'a new and properly practical dimension' of subjectivity (OSA 101/121). The ontology of the flesh is a presupposition for this practical dimension, since it outlines the pragmatic dimension of action. The subject is capable to act through its body. The subject and the world comes together in the body, since the "I can" depends on the body that is, at the same time, *my* body

and *a* body among the other bodies in the world (OSA 111/135). However, the subject is more and other than its capacities to act. The subject is its actions. Among the possible options that are made possible by its capacities, the subject decides on a certain option by realizing it through action. Now, these actions may be 'blamable or praiseworthy' according to certain norms or standards (these will be dealt with in the concluding section), in short, they are characterized by their imputability. In this way, the agent and the action are related by means of the concept of imputability. The action refers to a self that has chosen to act in one way and not in another, and thereby the self puts itself at stake in the actions that it performs. Thus, the subjectivity of action is ultimately a question of the self that initiates the action: 'with imputability the notion of a capable subject reaches its highest meaning, and the form of self-designation it implies includes and in a way recapitulates the preceding forms of self-reference' (CR 106/158). Said differently, the difference between action and event is that an action is personal whereas an event is impersonal. And this difference reveals the difference between motive and cause. A motive *initiates* an action when chosen and realized by a subject whereas a cause *continues* a larger chain of reactions without an explicit initiation, which therefore cannot be qualified as an action but only as an event. And since the subject somehow initiates the action, that is, he or she is the possessor of the action, attention must necessarily shift from the action to the agent. Whereas we cannot place a responsibility concerning events (since they are anonymous), actions are always being evaluated in that they originate in a subject that could have done otherwise. So, in order to clarify the personal aspect of action, Ricoeur turns to the practical dimension of subjectivity. His analyses therefore shifts from the logical and pragmatic aspects of action to the ethical and moral aspects under the guidance of the concept of imputability: 'Imputability, we shall say, is the ascription of action to its agent, *under the condition of ethical and moral predicates*, which characterize the action as good, just, conforming to duty, done out of duty, and, finally, being the wisest in the case of conflictual situations' (OSA 292/338).

However, before dealing explicitly with the practical dimension of subjectivity, Ricoeur develops the concept of narrative identity that he claims will function as a bridge between the descriptive and normative aspect of subjectivity, namely by relating 'describing, narrating, prescribing' (OSA 140/166). The attempt to relate the descriptive and the normative aspect of subjectivity is an important feature of his theory, because it instantiates his idea of an integral notion of subjectivity. The notion of subjectivity is neither purely descriptive nor purely normative. It rests on an ontology capable of explaining a being that is rooted in both a factual and normative dimension. On the one

hand, the subject is a part of physical nature, which means that we must pay attention to the embodied nature of subjectivity (ontology of the flesh). On the other, the subject is a member of humanity, that is, it lives together with other human subjects who demand that the subject considers its existence a co-existence. Moreover, it is situated in a certain social and cultural context, inscribed in a tradition, and governed by specific rules and norms that cannot be reduced to the laws of physical nature. These two heterogeneous dimensions pertain to different ontologies. One that is grounded in the causality of anonymous events, and another that considers the normative nature of subjectivity: 'Are these two ontologies mutually exclusive? I do not think so; they are, in my opinion, simply different by reason of their starting points, which themselves cannot be compared' (OSA 86/107). Although Ricoeur focuses on a normative ontology, he is well aware of the possible integration of a normative ontology with a causal one.

Now, however, I shall take a closer look on the concept of narrative identity in order to clarify how narrating may bridge the apparent gap between description and prescribing.

The Narrating Subject: Personhood and Identity

The difference between cause and motive turned the analyses of action away from the action to the agent who performs the action. The agent is not just a cause among other causes in the anonymous chain of world-event. On the contrary, the agent is a self other than a part of the world. Now, a descriptive approach to human action must concentrate on *how* the action inscribes itself into the workings of the world, that is, *how* does the subject cause a change in the world; or, is it legitimate to say that the subject, autonomously, enacts an effective change in the world? Is the action not only part of a larger-scale system of anonymous cause and effect? Said differently, the descriptive analyses of action concentrate on the "why" and the "what" of a certain action. On the other hand, the normative approach turns to the agent and emphasizes the "who" of the action[16]. In this sense, the normative approach exceeds that of a theory of action because it surmises that personal action must include normative features: 'An agent is not *in* the far distant consequences as he or she is in a sense *in* his or her immediate act. The problem is then to delimit the sphere of events for which the agent can be held responsible' (OSA 106/130). So, we are left with two distinct approaches to human action that are seemingly incongruent. However, as we have seen, Ricoeur does not consider

[16] The terms descriptive and normative are both descriptive in the sense that they are both value-free accounts. The difference is that the one (the descriptive) focuses on causal relation, whereas the other (the normative) deals with the values and norms of the "who" that did the action.

the two approaches exclusive of one another. On the contrary, he attempts to connect the two, and in order to do so he employs the concept of narrative identity.

The concept of narrative is supposed to bridge the factual and the normative dimension of subjectivity by introducing the temporal dimension of existence, namely the fact that the subject has its own history (OSA 113/137; CR 101/153). So, although the fact that the subject has its own history might seem intuitive and trivial, Ricoeur develops his conception of personal action on the problems concerning personal identity by emphasizing the historicity of human existence (OSA 114n/138n). How does the subject remain identical to itself throughout its existence in time?

We saw earlier that Ricoeur uses the concept of narrative identity in order to cope with the aporetics of human temporality. On the one hand, the subject is firmly rooted in physical (cosmological) time, which means that the subject is bound to change together with everything else in the world. On the other, the subject experiences time somewhat differently than merely the changing of seasons; an hour can sometimes seem to last a year and other times it vanishes in a second. This is what Ricoeur calls phenomenological time. Both conceptions of time play an important role in understanding subjectivity and cannot be reduced to one another. Therefore, Ricoeur proposes a third conception of time, namely time as configured by narratives. We shall now look at how the question of personal identity may be approached by means of narrative configurations of time.

I answer the question in two moves. First, I show how narratives articulate the difficulties concerning personal identity by stressing the dynamics involved in the constitution of identity: How can the person be said to remain the same person through the changes it makes and suffers over time? Secondly, I will clarify how Ricoeur uses the notion of narrative configurations to emphasize the importance of the concept of responsibility to personal identity.

Narrated time is supposed to include both the passive and active aspect of human temporality in that it pays due attention to how time works on the subject as well as to how the subject copes with time. In difference to a plant or a dog, a human subject does not remain passively receptive to the workings of time. For example, we go to a hairdresser when our hair becomes too long, stop eating cholesterol rich food when we become older or plan our future with respect to our past. We adapt to time, reflect, and work on time. Therefore, our being this particular subject is thoroughly determined by time, because time poses a challenge to our identity in the sense that it is ‘a factor of dissemblance, of divergence, of difference’ that we, as persons, have to deal with (OSA 117/142).

Ricoeur individuates three components in personal identity: 1) numeric identity, 2) qualitative identity, and 3) uninterrupted continuity, which on their part express the main principle of permanence in time (OSA 116-7/140-1). A person must fulfill all three criteria over an unspecified period of time in order to be said to be the same person and not two different ones. Numeric identity refers to the fact that a person must remain the same physical object through all kinds of imaginable changes in time. A person might loose all her physical characteristics over the years or completely loose her memory of all deeds done or events occurred to her and still be the same person in a numerical sense, because she is still the same physical matter that has evolved through time, as for example, verified by looking at her DNA-code (OSA 117/142). Nevertheless, we might find it difficult to accept that her *personal* identity has remained the same if all the characteristics that defined her as a certain person at some point has changed into something completely different. Therefore, to understand personal identity we must also pay attention to the qualitative and the uninterrupted continual aspect of the person in question. We identify a person according to certain character traits. She walks in a certain way, her voice has that specific intonation, her temper is bad late in the afternoon, and she loves fishing and is faithful to her longstanding principles about monthly charity to the less fortunate, the discipline of hard work and early rising, and the necessity of death sentence. Now, time plays an even more critical role with respect to this kind of identity (CR 65-6/101). As we age we remain the same physical matter but our dispositions and character traits may change, and that sometimes drastically. Could we say that a person is still the same person if his or hers principles, ideas and overall life-plan are changed beyond recognition?

The question is extremely difficult. For example, we allow that a cold-blooded and horrible murderer can change, serve her sentence and become a better person that has nothing whatsoever in common with the crimes done in the past. She is quite literally a different person. Thus, we judge a person on how he or she acts now and not on past mischiefs or benevolence for that sake. On the other hand, we are well aware that it is, in fact, not another person. It is still the murderer that has changed into something better because she has learned from her terrible deeds, transformed her principles and worked out a significantly different life-plan for herself. How, faced with these difficulties, is it possible to speak about the permanence of personal identity through the sometimes dramatic changes in time? Ricoeur is convinced that '[t]he idea of structure, opposed to that of event, replies to this criterion [permanence] of identity, the strongest one that can be applied' (OSA 117/142). We must concentrate on the structural permanence of the person and not on the events that seem to dissolve personhood into fragmented pieces of a self. In other words, we must find

what remains invariable in the different relations in which the person finds itself through the course of time: 'The entire problematic of personal identity will revolve around this search for a relational invariant, giving it the strong signification of permanence in time [...] a form of permanence in time which can be connected to the question "who?" inasmuch as it is irreducible to any question of "what?" A form of permanence in time that is a reply to the question "Who am I?" (OSA 118/142-3. Translation slightly modified).

Ricoeur now attempts to find this 'relational invariant' of personhood by analyzing identity with reference to two opposite models of permanence in time: *character* and *keeping one's word* (idem). These two models of permanence form a dialectics that, articulated by means of a narrative structure, might capture and explain the subtle difficulties surrounding the notion of personal identity. Ricoeur puts it like this: 'My hypothesis is that the polarity of these two models of permanence with respect to persons results from the fact that permanence of character expresses the almost complete mutual overlapping of the problematic of *idem* and *ipse*, while faithfulness to oneself in keeping one's word marks the extreme gap between the permanence of the self and that of the same and so attests fully to the irreducibility of the two problematics one to the other […] the polarity I am going to examine suggests an intervention of narrative identity in the conceptual constitution of personal identity in the manner of a specific mediator between the pole of character, where *idem* and *ipse* tend to coincide, and the pole of self-maintenance, where selfhood frees itself from sameness' (OSA 118-9/143).

In order to understand this program, we first have to briefly clarify the terms that he is introducing.

The notion of *character* was introduced in the previous chapter as the person's immutable perspective and opening to world, the so-called 'existential difference or negation', which we cannot change (our birth place, physical constitution, and, to a certain degree, also our habits). Ricoeur remains faithful to his early analyses, but now he concentrates on the temporal dimension of character, which brings into question the immutable status of character. Some aspects of our character loose their feature of immutability when considered in relation to time in the sense that our habits and dispositions may change through the span of our life. So, '[c]haracter, I would say today, designates the set of lasting dispositions by which a person is recognized' (OSA 121/146). In this way, the notion of character functions as a criterion by which we are able to identity a person in relation to the general things or traits that we can describe, without asking about the self behind the different predicates. We can therefore say that '[c]haracter is truly the "what" of the "who"' (OSA 122/147).

On the contrary, the model of *keeping one's word* 'expresses a *self-constancy* which cannot be inscribed, as character was, within the dimension of something in general but solely within the dimension of "who?"' (OSA 123/148). This kind of permanence in time captures the self that we cannot identify by means of mere character traits, that is, the self that is more than the sum of its descriptively recognizable traits. To keep one's word is to defy the changes that time brings about. Ricoeur exemplifies this dynamics with the nature of promise. Say that I have promised a friend to visit him in Berlin within a few months. However, work piles up, money runs low and my desire to visit him seems to subside with respect to other and more interesting agendas. Nonetheless, I have given him my word that I will be coming, and not to do so would therefore be inconsistent with what friendship stands for, namely trust and sincerity. It is obvious that this kind of permanence in time deals with another aspect of identity that cannot be accounted for in neutral, descriptive terms but involves a justification able 'to respond to the trust that the other places in my faithfulness. This ethical justification, considered as such, develops its own temporal implications, namely a modality of permanence in time capable of standing as the polar opposite to the permanence of character' (OSA 124/149).

Thus, Ricoeur individuates two distinct poles of personal identity: A descriptive form which he names identity of the same (idem) and a normative form which is named identity of the self (ipse).

The first expresses the passive form of personhood that remains the same in that it fulfills the three criteria for identity through time (numeric, qualitative, and uninterrupted continuity) without any significant activity on behalf of the person itself. The second demands that the person is constantly aware of how to remain faithful to his or her promises and engagements in light of certain values, norms and social codices. On the one side (identity of the same, i.e., character), we have that which is seemingly the most stable aspect of personhood, namely the physical characteristics and longstanding dispositions shaped by upbringing, education, and convictions, which are all gathered in the notion of character. On the other (identity of the self, i.e., keeping one's word), we have the more dynamic and impalpable normative aspect, which has to do with how the person acts according to the norms and procedures that govern the social interface between him and other persons.

Once clarified the key notions (character, same, idem vs. keeping one's word, self, ipse), we can begin to understand how narrative identity may be a help in approaching the difficulties concerning personal identity.

Although character is the most stable form of identity, i.e., that by which we are re-identified and thereby recognized, many of our character traits are still a result of ongoing conscious choices made during our life-time; choices become habits and long-term dispositions, which turn into something that we do no longer choose; or said more eloquently, habits are 'the return from freedom to nature' (OSA 121/147). The formation of habits and seemingly unconscious or automatic dispositions to act is our most obvious and docile tool to cope with alterity, since they internalize the alterity outside of us (the world and the other) and make it a part of our person (OSA 122/147). Our dispositions to act are shaped by how we are brought up and by our social, geographical, and cultural context. If Ihave grown up in a violent environment, my first reaction to fear and provocation is more likely to be an act of violence than if I had always been taught to avoid the use of violence at any cost. Thus, a major part of our personhood is determined by how our habits and dispositions are developed. A person is, however, more than its habits and dispositions. It is a self that can relate itself to the nature of its character, since it can choose to follow its inclinations and dispositions or choose not to. The identity of the person depends on the dialectic of self and character; or as Ricoeur writes: 'The dialectic of innovation and sedimentation, underlying the acquisition of a habit, and the equally rich dialectic of otherness and internalization, underlying the process of identification, are there to remind us that character has a history which it has contracted, one might say, in the twofold sense of the word "contraction": abbreviation and affection', and he goes on to explain what narrative identity might contribute to the understanding of that dialectic, namely 'to balance, on one side, the immutable traits which this owes to the anchoring of the history of a life in a character and, on the other, those traits which tend to separate the identity of the self from the sameness of character' (OSA 122-3/147-8).

The key-notion here is 'the history of a life'. Narratives help evidence the structure and dynamics of our identity in that they articulate the reasons for our character and dispositions. The narrative defies, so to say, the immutability of the character, because it makes the otherwise inalterable part of our personhood (the same or idem) less fixed and a dynamic part of our personal history. Through narratives, we may become aware of how we came to have such a strong character or such violent dispositions, and not just accept that, in the end, that is just how we are. The dialectic of selfhood and sameness has a history. We are who we are because of the history of actions and sufferings. To approach our identity as a narrative structure makes us capable of distinguishing between what we have done ourselves and what has not originated in ourselves but produced by others factors such as upbringing, friends, birthplace, education etc.. We are always the same

person, but our personhood is characterized by a tension between what we do and what we are. I might be a kind and loving person, but in order to remain so I must continue to act according to that which makes me a kind of person. Or, I might be a cruel and violent person, but that does not mean that I am destined to remain such a person. Now, narratives articulate this dynamics because '[w]hat sedimentation has contracted, narration can redeploy' (OSA 122/148).

At this point, it might be clear that Ricoeur continues to elaborate the structural schemes that he developed in the early works. The subject is marked by a fundamental non-coincidence, sometimes even a conflict, in the core of its being, namely between vital and spiritual desires where the former tends toward pleasure according to the immediate well-being of the subject itself whereas the latter seeks a more total form of happiness that must include the presence of and co-existence with the other subjects. This problematic is now transferred to the practical level. How can the subject be said to remain an integrate person when continually faced with the threat of disintegration and fragmentation that temporal (co-)existence with other persons brings about in the course of a life?

Once again, we are faced with the problem of conflict and non-coincidence. This time, however, the experiencing, acting, and feeling subject is considered not as a structural being but as a concrete person situated *and* acting in a certain historical, temporal, and socio-cultural context. And as the foregoing section has pointed out, the question of personhood is closely linked to that of action, since human action is personal action in that the subject always initiates the action according to its own motives such as desires, inclinations and rational cognitions, in short, according to its own hierarchy of values (i.e. the configuration of the practical space of reason; this will be dealt with in the last section). Responsibility is what renders human action distinct from the realm of mere, anonymous events. And yet, responsibility is a problem, not a fact. The issue at stake here is to narrow down on where responsibility begins. We have seen that narratives may help individuate the "who" in the "what", that is, to what extend can we be said to be responsible for what we are. For example, how are we to blame for our violence, if that is all that we have ever experienced? Of course, narrative identity cannot solve the difficulties involved in personal responsibility, but it can bring forth nuances to the problem. Personal identity is not just a question of what we are or how we remain the same, but also how we relate ourselves to the person that we are. The question is then how are narratives supposed to clarify our understanding of responsibility in relation to personal identity?

Understanding personal identity becomes a very difficult task, when we want to account for identity in other terms than numeric identity. Qualitative identity, or the 'relational invariant', seems

to be impossible to grasp, since the person keeps changing throughout his or hers entire life. Ricoeur, however, chooses to confront this problem with the concept of responsibility. And, according to his reflections on motive and cause, responsibility is originally a question of imputability: 'It is for the other who is in my charge that I am responsible. This expansion makes what is vulnerable or fragile, as an entity assigned to the agent's care, the ultimate object of his responsibility' (CR 108-9/162). My identity as a person is ultimately a question of my actions in regard to the other. This is where the fragility of my personhood becomes clear. If my identity only depended on me, then I could change my opinions, motives, and principles at my heart's desire without ending up in difficulties. I did not have to account for my actions to anyone. Now, this is not the case. I always act as a person in the eyes of the other. I would be unrecognizable, as a person, if I one day supported capital punishment and the next day marched for the abolition of this legal action, or said on Monday that I love cats and hate dogs, and then on Wednesday claimed that I hate cats but love dogs. In order to maintain personal integrity, I have to remain identical in at least some matters, that is, I am responsible for what I do and say. Now, the problems arise when we have to determine how much and in what matters can a person change and still remain the same 'relational invariant', the same person? This is where narrative theory comes in. There seems to be a gap between describing a person and prescribing what a person must do to remain a person, i.e., between the ascription of an action to an agent (theory of action) and the imputation to the agent of an obligation to act in a certain way in order to remain the same person (ethical theory) (OSA 152/180). The key-concept that is supposed to bridge this gap is 'narrative configuration' (OSA 142/169). To consider the structure of action in terms of narratives is a way to extend 'the field of practice' so that the agent's various actions become configured according to a hierarchy of practical units that calls for a revision in the concept of action (OSA 152-3/181). The practical field is extended and qualified, because narratives view actions from a top-down perspective that differentiates human events from neutral occurrences and furthermore consider actions over the course of time (OSA 141/169). Or with Ricoeur's words, 'Telling a story is saying who did what, why and how, by spreading out in time the connection between these various viewpoints' (OSA 146/174. Translation modified). Contrary to a causal explanation of action, narrative configuration is concerned about the qualitative aspects of personhood that become articulated in the story. Whereas a causal model tends to dissociate the agent from the normative aspect of its action, the narrative model emphasizes this aspect because no story is 'ethically neutral' and therefore 'proves to be the *first laboratory of moral judgment*' (OSA 140/167). Thus, narrative configuration provides

'*supports* and *anticipations*' (OSA 152/180) for an ethical theory. In using the structure of literary fiction as a model for personal identity, Ricoeur asserts that a story always involves agents and sufferers and opens up a field of evaluations where we judge the value of an action (and thereby its agent) in relation to how it affects another person who thus becomes a sufferer of the action done (OSA 144/172). In this way, the narrative configuration articulates and emphasizes what is actually the case with all human action, namely that 'every action has its agents and its patients' (OSA 157/186). Of course, this is not to be taken literally as, for example, when I pour myself a cup of coffee, I then determine the suffering of another human being. However, our actions (even the most insignificant one) are, somehow, always inscribed in a larger complex of human relations. This is what is emphasized in the narrative configuration, and which tends to be confused and hidden in our everyday experience (OSA 148/176), and therefore '[i]t is precisely because of the elusive character of real life that we need the help of fiction to organize life retrospectively, after the fact, prepared to take as provisional and open to revision any figure of emplotment borrowed from fiction or from history' (OSA 162/191-2).

Narratives are an attractive model for explaining personal identity because it does not exclude causal explanations (bottom-up perspective), but seeks to integrate them in a wider explanation. It is true that narrative configuration approaches the question of personal identity from a normative, top-down perspective, but it operates with a dialectics that emphasizes the complexity of human nature, namely that a person is both part of nature and humanity. We cannot understand personal identity in isolation from one of these aspects. The person is determined both by its physiological constitution and by its being part of human interrelations: 'The practical field then appears to be subjected a twofold principle of determination by which it resembles the hermeneutical comprehension of a text through the exchange between the whole and the part. Nothing is more propitious for narrative configuration than this play of double determination' (OSA 159/187).

This play of double determination is emphasized in the narrative configuration of our practical field of experience because, although narratives are structured around what Ricoeur calls '*imaginative variations*' of our personal identity (OSA 148/176), and that these always involve 'an unstable mixture of fabulation and actual experience' (OSA 162/191), our personal identity remains grounded in 'the existential invariant of corporality and worldliness' (OSA 151/179). Imaginative variations are instable because they deal with the more impalpable features of personal identity: ideas, wishes, dreams, norms, social heritage, upbringing etc. and are therefore liable to fabulation, self-deception and denial. This is, nevertheless, balanced by the invariable nature of corporality and

worldliness in the sense that the stories we (and others) make about ourselves find their verification or appropriateness in relation to our body and to the other persons around us. Our experience of ourselves and others is configured by narratives, but the narratives themselves are always grounded in the nonnarrative dimensions of affectivity and the presence of the other. The previous section dealt with affectivity and the ambivalence of the body as being, at the same time, mine and just another object in nature. The narrative configurations of our self-experience are always determined by this ambivalence, since we may ascribe to ourselves physical qualities that we, in reality, do not possess, but this self-deception will, normally, shatter confronted with the mirror of alterity (my actual body, the world, and the other). Even though Ricoeur focuses on the normative aspect of personal identity, his model is, as we have seen, compatible with a more descriptive, causal approach.

However, it is the normative aspect that Ricoeur focuses on. Narrative configuration is supposed to bridge the descriptive and prescriptive level of personal identity, which means that it is a necessary but not sufficient explanation of personal identity. On the one side, it is conditioned by the descriptive explanation of the body, and on the other, it is determined by the concept of responsibility. Now, the concept of responsibility is emphasized by the narrative configuration, but it also transcends the explanatory power of narratives, since they may clarify the dynamics of personhood but leave the problem of responsibility open: 'The story of a life includes interactions with others […] In the test of confronting others, whether an individual or a collectivity, narrative identity reveals it fragility' (CR 103-4/155-6).

Responsibility is one of the key-concepts that both the schematic notion and the redressed notion of subjectivity have revolved around. Ricoeur has considered it a fundamental concept in the constitution of human subjectivity since his early structural analyses in FN (FN 56-7/55). Therefore, in these concluding pages, we must find out how this concept can tie together the results that the preceding analyses have brought forth. In order to do so, we have to take a final look on the notion of a practical space of reason.

Ethical Experience: The Practical Space of Reason

In this last section, I shall argue for what I have named the practical space of reason. By and large, I will try to show that subjective experience is configured as a practical space of reason, and, furthermore, that this is a central idea in Ricoeur's theory of subjectivity. To say that subjective experience is configured as a practical space of reason means that human subjectivity must be

approached and understood from the perspective of such an experiential space of values and norms. However, this is not the same as to say that the human subject cannot be approached from other perspectives such as physics and biology (i.e., the empirical sciences), which we have already seen regarding the ambivalence of the body; only that such approaches has to be integrated with an account of the subjective experience of the world, the other and the self in order to give a cogent account of human subjectivity.

The section is divided in three parts. First, I will clarify what I mean by the configuration of experience according to a hierarchy of heterogeneous values by analyzing further the concepts of value and desire in relation to human practices. Here we find that narrative configurations, as a way of explaining human action, are a preliminary exploration of the practical space of reason. Secondly, I shall explain the relation between the self and the other in the configuration of experience. This relation shapes our action in the world and is critical for the way we experience ourselves as a self and recognize the other subject as a self as well. Finally, in way of conclusion, I argue for the existence of a primary ethical experience. Our experience of the world is primarily ethical in the sense that we perceive objects and especially other subjects in world according to some sort of hierarchy of values. Our perception is basically saturated with normativity. As we develop into adult subjects, become aware of ourselves as an individual person with a specific history and embedded in a certain society with particular norms and values, this normativity enriches and grows more fine-grained and complex. I claim that responsibility is the fundamental notion for way we understand ourselves as persons. Identity and self-esteem both find their correlate in our responsibility toward the other person. Thus, I arrive at the concept of responsibility as the key-concept that binds together the various aspects of Ricoeur's notion of subjectivity in the development of an ontology that account for the subject as a person as well as a perishable part of physical nature. I claim that the normative experience of the world is primarily ethical, because our existence in the world is a coexistence with other subjects, and that this coexistence affects our experience in a radical way. The legitimacy of my possible actions in front of the other subjects determines the way I experience the world as a space for the attestation of my selfhood.

I shall not deal with the specific nature of what Ricoeur calls his 'little ethics' (OSA 290/337), since the primary interest of the present investigation is about the relation between subjectivity and ethics, and not about which kind of ethical theory is the most appropriate one. Ricoeur's theory of subjectivity is an attempt to clarify the structures and dynamics that are at play in human subjectivity and, further, an argument for a softening of Hume's watershed between judgments of

fact and judgments value, that is, between 'is' and 'ought' (OSA 171/201-2). In fact, Ricoeur is not particularly interested in a normative ethical theory, but concerned with metaethical questions such as the ontological and epistemological nature of human values and the relation between human nature and ethical dispositions (i.e. imputability, obligation, and responsibility). He delineates a human ontology that articulates the intimate relation between facts and values in order to answer how we can legitimate a claim for ethics in a seemly indifferent physical world. But I am getting ahead of myself. Let us take the argument step by step and start by first looking at the configuration of experience.

The Configuration of Experience

In the previous chapter, I defined the practical space of reason as experience of the world configured as a hierarchy of heterogeneous values generated by the self, the world and the other. These values are originally produced in the affective dimension of the subject that experiences itself as a self with certain desires. The manifold of subjective desires was structurally divided in two distinct kinds, namely vital and spiritual desires. These generate a non-coincidence, often a conflict, in the heart of the self between the self as part of physical nature that seeks its own well-being and the self as part of a totality, that is, humanity. This basic conflict is the ontological origin of the fragility and unrest of the human self. Now, our experience of the world is shaped by our interaction with the objects in the world. The concept of intentionality made it clear that we are not passive spectators, but active participants in the unfolding events of the world. Objects and events do not just appear before us, but are loaded with meaning that is relevant to us as existing in the world. Objects gain value through the inverse movement of intentionality (from the objects to the subject) revealed by the feelings that we experience. These feelings manifest themselves as values (sometimes vague and diffuse and sometimes clear and strong) according to the rough distinction of vital and spiritual desires. This distinction is important to keep in mind when we in the following look at how values are generated in the subject that is embedded and situated in a concrete, historical world.

He develops his hermeneutics on the notion of conflict produced by the fundamental tension between vital and spiritual desires. The originating affirmation is now changed to attestation but the meaning remains the same: 'attestation is fundamentally attestation *of* self […] a trust in the power to say, in the power to do […] attestation can be defined as the *assurance of being oneself acting and suffering* (OSA 22/ 34-5). However, this assurance of being oneself is complicated by the fact

that the subject does not act and suffer in timeless isolation, but is always situated in a certain context determined by the world and the other. The attestation of the self is therefore conditioned by the existence of the world and the being of other selves. Our power to say and do is conditioned by what Ricoeur in his early writings called 'the existential difference', that is, that we live in a world inhabited by selves whose existence is different from ours. This means that we cannot just say or do anything in order to attest ourselves; if we want to affirm ourselves as a self, we must do so by recognizing the existence of other selves that are part of the world that we all inhabit: 'the exploration of human capacities should join to each modality of the "I can" an often tacit correlation between self-assertion and some reference to others. Self-assertion does not signify solipsism' (CR 252/363).

What does this rejection of solipsism mean for our rough distinction of vital and spiritual desires? The subject's self-assertion becomes problematic. On the one hand, it wants to assert itself as a self at any cost. It has specific desires, wishes, ideas, and plans for the future that it feels that it must realize in order to be the self that it wants to be. On the other, it cannot disregard (if it wants to exist in the same world as other subjects) the desires, wishes and plans that the other subjects have and want to realize. Therefore, if the actions of the subject shall work toward the attestation of the self, the values that orient its desires must involve both these aspects of self-assertion. Thus, once again, we have to explore the basic idea of a hierarchy of values but now in relation to the embedded and situated nature of the subject.

We saw in the structural analyses that values are primarily rooted in the spontaneity of the body and originates as certain *needs*: 'Through need, values *emerge* without my having posited them in my act-generating role: bread is good, wine is good' (FN 94/90). However, the fact that the subject is embedded in a socio-cultural context and situated among other subjects demands that this basic idea of values is developed into more complex notions of good and bad. A thing or an act is not necessarily good because it satisfies my immediate bodily need. It must also satisfy my more general idea of well-being. I do not live only to survive; neither are my needs limited to bodily needs nor my desires restricted to bodily satisfaction. My values may originate in the biological function of needs (a sort of homeostatic balance), but in fully developed adults they are also defined by a more reflective instance of well-being that Ricoeur names (inspired by the moral philosopher Alasdair MacIntyre) 'the good life': 'Life must be gathered together if it is to be placed within the intention of genuine life. If my life cannot be grasped as a singular totality, I could never hope it to be successful, complete' (OSA 160/190). Thus, we see how the structural distinction between vital

(immediate) and spiritual (totality) desires, as the core of human affectivity, remains a fundament for the hermeneutic analyses; or as Ricoeur puts it: 'if there is a being of the self – in other words, if an ontology of selfhood is possible – this is in conjunction with a *ground* from which the self can be said to be acting' (OSA 308/357; translation slightly modified). That ground is precisely the affective dimension of human experience that was analyzed in the last section of the previous chapter.

What is important, at this point of the theory, is to clarify how the notion of 'the good life' affects and develops the idea of a structural hierarchy of heterogeneous values. To do so, Ricoeur employs the analyses he made concerning narrative identity.

The concept of value grows more complex and refined when the subject is conceived as a self embedded in a socio-cultural context and situated among other subjects. This complexity is due to the fact that the concept of value is always a compromise, and therefore only qualifies as a quasi-concept: 'In this sense, I hold the quasi-concept of value to be a compromise term, at the point of intersection of the claim to universality and the admission of the historicity of certain derivative duties to which corresponds the right of others to make claims on us' (OSA 289/336. Translation slightly modified). Value is a quasi-concept because it is a variable and heterogeneous entity that does not allow for a clear definition. Values cannot be defined a priori, but are always shaped by a certain context and a specific situation. Nonetheless, some values still seem to possess some kind of universal aspiration that separates them from mere whims and preferences and procures them a more firm status among our motives for a given action. This is where narratives come into the picture. The structure of narratives may help clarify the complex notion of values by individuating the different aspects of situation, culture, and universality involved in certain values.

The subject gains its singularity through the temporal unity of its life in the sense that the self cannot be separated from the actions and sufferings that constitute its identity through the changes in and of time (OSA 147-8/175). This is exactly what is emphasized in the narrative configuration of experience. Narratives ascribe actions to an agent and explore the implications and effects of those actions, especially with respect to interpersonal relationships: 'stories are about agents and sufferers' (OSA 144/172). We saw how personal identity can be understood in terms of narratives, because our identity as a person depends (partially) on stories that explain the *why*, the *what*, and the *how* of our being this individual self with certain character traits, dispositions, idea(l)s, plans, and behavior. By providing answers to these questions, narratives organize our confusion concerning personal identity. Now, narrative configurations not only clarify questions regarding

personal identity, but they also assist us in our explanation of human action and behavior. Narratives structures extend the practical field of human action by configuring the otherwise incomprehensible multitude of actions into 'a hierarchy of units of praxis that, each on its own level, contain a specific principle of organization, integrating a variety of logical connections' (OSA 153/181; translation modified). An action is understood according to these units of practices, since values, motives, and decisions find their explanation in light of the logic governing a certain practice. For example, I move my arm and finger *because* I have to switch on the light *because* I need to see in an otherwise dark room *because* I have to find the refrigerator *because* I must find the sandwich *because* I am hungry. Each of these *because* signifies a practical unit that explains why I do a certain action. They run from the most basic action, such a moving my finger, to a more global level that involves actions concerning such serious issues as life projects or general ideas of happiness, goodness, and justice. Any approach to the dynamics of this hierarchy of practices must begin with 'the simple fact that the practical field is not constituted from the ground up, starting from the simplest and moving to the more elaborate constructions; rather it is formed in accordance with the twofold of ascending complexification starting from basic actions and from practices, and of descending specification starting from the vague and mobile horizon of ideals and projects in light of which a human life apprehends itself in its oneness' (OSA 158/187).

Thus, the logic of an action can be explained on different levels in accordance with the unity of practice we take into consideration, and the different levels possess their own kind of logic. I move my finger to grab, I grab to pull out a chair, I pull out a chair to sit, I sit because I am tired etc. Nonetheless, the logics involved at different levels are somehow interrelated. I may perform innumerable basic actions without paying attention to my horizon of ideals and projects, but often even simple actions involve the presence of other persons. Thus, my action becomes an interaction, and we return to the fact that 'every action has its agents and sufferers' (OSA 157/186). When I perform an action, that action is always saturated with normativity because it has a value both to myself and other persons. In this way, Ricoeur argues that our understanding of human action is always oriented according to a hierarchy of heterogeneous values constituted by rules that point in the direction of moral rules (OSA 155/183). However, it is important to emphasize that these rules are first of all rules of meaning, as for example the rules of chess: 'The rule, all by itself, gives the gesture its meaning: moving a pawn; the meaning stems from the rule as soon as the rule is constitutive, and it is so because it constitutes meaning, "counting as". The notion of constitutive

rule can be extended from the example of games to other practices, for the simple reason that that games are excellent practical models' (OSA 154/183).

The rules that constitute the meaning of action are governed by the configuration of experience as a practical space of reason, where we experience the external alterity as an organized whole, a world that means something to us as individual subjects. Actions and events present themselves in terms of values according to the practice that they are inscribed in. When the subject wants to assert itself as a self in the world, the practices by which it realizes this assertion have to pay attention to rules that govern the practice. Values are defined by the rules that constitute the practice. To put it simple, something is good if it attains the goal of the practice and bad if it hinders or prevents the fulfillment of the practice. For example, if I successfully grab a chair that action is good in light of the practice of getting a chair; or opening the refrigerator to get food; or working to get money. Thus, our experience of the world is organized by a hierarchy of values pertaining objects, things, persons, and actions. When we perceive objects they present themselves as meanings according to different practices and their value is determined by the fulfillment or failure of the practice. Subjective experience is primarily configured according to these practices. Human subjects do not experience a fragmented succession of indifferent objects, but things and persons inscribed in the, for the most part, coherent unity that we experience as our existence in the world. This human existence can be approached as a narrative unity that yields our experience its character of totality, a life. Ricoeur puts it like this: 'The term "life" that figures three times in the expressions "life plan," "narrative unity," and "good life" denotes both the biological rootedness of life and the unity of the human being as a whole, as that human being casts upon itself the gaze of appraisal […] As for the term "narrative unity," the aspect that we are emphasizing here is less the function of assembling-together, performed by the narrative at the summit of the scale of praxis, than the connection the narrative makes between estimations applied to actions and the evaluations of the characters themselves' (OSA 178/209).

This has an important impact on the rudimentary notion of desire that was developed in the last chapter. We saw that human action is ultimately rooted in two basic kinds of desire, vital and spiritual desire. This schematic notion of desire was the basis of human affectivity that was revealed by our feelings. These generated a non-coincidence in the heart of the subject that was felt as a conflict between a desire to self-preservation (my own well-being) at every cost and a desire for happiness as a kind of totality that procured global meaning to the manifold of different actions done by the subject. I noticed that the conflict could not be solved at a reflective level, but had to be

understood through the various crystallizations of these basic desires in the concrete existence of the subject, i.e., embedded and situated in language, society, culture, history, economics, and so on. Thus, we have to understand the quasi-concept of values on the basis of these fundamental desires, but we need to approach them in term of the rules that govern the practices of a subject embedded in a certain socio-cultural context. Our hierarchy of values is therefore determined by how we assert ourselves through practices in a particular society governed by certain rules and norms of behavior. Values are partially contingent because of this dependency on context, but, at the same time, at least some values seem to aspire to a kind of universality. These universal values are for the most part those involved in interpersonal relation. Most actions involve and influence the presence of other persons; or as Ricoeur puts it: 'Practices are based on actions in which an agent takes into account, as a matter of principle, the actions of others' (OSA 155/184). The presence of other subjects has critical impact on our hierarchy of values (and thus the configuration of experience), because values regarding the presence of the other claim a greater degree of universality than other values.

To understand why this is so, we have to analyze what the presence of the other subjects means to the self, and in this connection we need to take a closer look on the notion 'the good life'.

The Self and Other Selves: 'The Good Life'

The practical space of reason is thus configured according to certain values derived from constitutive rules of practices. I understand (i.e. they have meaning and therefore a value for me) actions and events in the framework of such practices: 'the self is essentially an opening onto the world, and its relation to the world is […] a relation of total concern: *everything* concerns me. And this concern indeed extends from being-alive to militant thinking, passing the way of praxis and living well' (OSA 314/363). Now, although everything concerns me, some things concern me more than others. My concerns originate in the basic desires, since my actions in the world are generated by my general desires for self-preservation and happiness. The world becomes the horizon for the realization of selfhood: 'The being of the self presupposes the totality of a world that is the horizon of its thinking, acting, feeling – in short, of its care […] There is no world without a self who finds itself in it and acts in it; there is no self without a world that is practicable in some fashion' (OSA 310-1/360). When I experience the world, I find it primarily configured as possibilities to satisfy these concerns. In the multitude of my concerns, the basic value of living is always counterbalanced by the more complex value of living *well*. I care about living, but just as much about living well. Now, the hierarchy of values that configures my experience of the world as a practical space of

reason is affected by my concerns, in particular my concern for living well. Therefore, we have to return to the notion of 'the good life' and see what this notion means to our conception of practices.

The notion of 'the good life' is a qualification of the existence of the subject conceived as a singular totality (OSA 160/190); that is, the subject considered as a unique self that seeks to assert itself through and, sometimes, in spite of time. All our actions, practices, and sufferings come together in the qualification of our existence as a totality. The basic desire of happiness (the totality of pleasure) imposes upon us the need to qualify our existence according to an idea of the good life. For many reasons, this is not a simple need to fulfill. How can we grasp our entire existence in a single notion such as 'the good life'? And what is actually meant by 'good'? As we have seen, the values by which we live our life are partially contingent and contextual which means that such a general notion as 'the good life' becomes extremely fragile (OSA 179/210). Nonetheless, Ricoeur is convinced about 'the idea of a higher finality which would never cease to be internal to human action' (idem). Human action is always interpreted according to a hermeneutical dialectic between the particular and totality, as it was pointed out by the narrative approach to personal identity (OSA 158/187); with a notion borrowed from the philosopher Charles Taylor, Ricoeur defines the human self as a 'self-interpreting animal' (OSA 179/211)[17]. On the one hand, our different practices have their own local value according to the logic determining their specific fulfillment. On the other, their value is influenced by the global idea of the good life, our existence qualified as a totality. Somehow, all our actions and practices are governed, consciously or unconsciously, by a higher finality according to this global idea. The obscure and intuitive nature of such an idea needs interpretation in order to become articulate. Therefore, although the specific idea of the good life may change in the course of our lives due to the mutable and contextual nature of our feelings and ideas and, though more stable, our character and dispositions, the general idea somehow remains. And as such it influences the configuration of our experience of the world, because it affects the values by which we live our life. We ascribe certain values to the objects that we perceive as things that matter to us; our concerns are determined by what we care about.

These concerns are experienced as values of more or less importance. Some are more contextual and vague than others; some arise and fade in the moment and depend heavily on haphazard circumstances whereas others seem to have a more universal claim on our concerns. As I have

[17] Ricoeur describes this dialectics in the following manner: 'between our aim of a 'good life' and our particular choices a sort of hermeneutical circle is traced by virtue of the back-to-forth motion between thde idea of the 'good life' and the most important decisions of our existence (career, loves, leisure, etc.). This can be linked to a text in which the whole and the part are to be understood each in terms of the other' (OSA 179/210).

already pointed out, the latter frequently involves the existences our own self and the existence of other selves. To understand the hierarchy of heterogeneous values that configure our experience, we must consider more closely the relation between the self and other selves.

As we saw in the last chapter, the subject is a self that continuously seeks to affirm or assert itself in the world. Its being is fundamentally characterized by what Spinoza called *conatus*: 'the effort to preserve in being, which forms the unity of man as of every individual' (OSA 316/366). This effort to preserve in being, to assert oneself in the world, is nevertheless conditioned by alterity, and in particular other selves. But why is the presence of other selves a condition for the assertion of our own self? Can we not just preserve our own being without regarding the presence of other selves? Ricoeur answers these questions by emphasizing the presence of alterity in the constitution of the self. The identity of the self is vulnerable and fragile because its needs cannot be satisfied by its own power and effort. Other than the original effort to preserve its own being, the being of the self is characterized by a 'need, hence a lack, that drives the self toward the other' (OSA 185/216). This need or lack originates in the spiritual desire for happiness that counterbalances the vital desire for self-preservation in the heart of the subject and expresses itself in our idea of 'the good life'. My identity as a self is qualified this idea, and therefore selfhood cannot be sustained in solitude, but requires the presence of other selves. I am who I am in the context of other selves, and my idea of 'the good life' is determined by the other selves with whom I am involved in my existence. Ricoeur employs two interrelated concepts in order to emphasize this strict relation between the being of the self and the being of the other, namely *self-esteem* and *solicitude* (OSA 180/212).

Self-esteem is 'the primordial reflexive moment of the aim of the good life' (OSA 188/220), since our life considered as a totality is valued in accordance to our own conduct. Our idea of 'the good life' finds its reflective correlate in the concept of self-esteem, because 'it is in appraising our actions that we appraise ourselves as being their author' (OSA 177/208). My self-esteem depends on the actions that I perform. The values that I ascribe to the unfolding of my life (actions as well as sufferings) are part of a more global evaluation of my existence. Now, the concept of self-esteem is inextricably linked with the concept of solicitude: 'To self-esteem, understood as a reflective moment of the wish for the "good life," solicitude adds essentially the dimension of *lack*, the fact that we *need* friends; as a reaction to the effect of solicitude on self-esteem, the self perceives itself as another among others' (OSA 192/225). The other considered as a self like me and yet autonomous with respect to me is constitutive for the idea of 'the good life'. My concern for the other does not originate only in a desire to dominate the other, but in a lack in my own existence

that generates a need for a dialogical coexistence with the other (OSA 180/212). We need the other not as a means to satisfy our own desires but as autonomous selves that constitute my global idea of 'the good life'. This need for others affects drastically the values by which I act in the world in the sense that the values that involve other selves are critical for my own existence. When the other selves address me 'in the second person, I feel I am implicated in the first person' (OSA 193/225). Therefore, my self-esteem, which is a dominant part of the global idea of 'the good life', depends on the presence of other selves that fulfill and satisfy the basic need of others. This need is felt as solicitude for the existence of the other. Selfhood is not constituted solely of mineness: 'To say self is not to say myself' (OSA 180/212). The other self is fundamental for my identity in the world, since the world is common world whose meaning is derived from the coexistence of different selves (OSA 332/384). My identity as a unique self depends on the other's recognition of my being so. The others recognize my actions as actions of a self with a certain history, ideas, and plans for the future; therefore, the unity and coherence of my various actions in respect to the idea of myself as a self is deeply influenced by the reception of these actions by other selves. I have a need to be recognized and vice versa: 'Becoming in this way fundamentally equivalent are the esteem of the *other as a oneself* and the esteem of *oneself as an other*' (OSA 194/226).

Hence, the presence of other selves is critical for my hierarchy of values because this hierarchy must reflect this recognition of the other as a oneself and oneself as an other. My actions and practices in the world is rooted in this primordial condition of selfhood, namely that the self is constituted by alterity in all its efforts to attest and affirm its own existence. The global idea of 'the good life' (happiness) involves this alterity and requires careful attention on part of the self to be realized. Therefore, the values that involve our relation to other selves have a somewhat more universally claim on us than other values in our existence. The solicitude for the well-being of the other defies the mutable and contextual nature of my other values. I cannot disregard the other in my actions and practices, because my identity as a person depends on the presence of him or her. Our experience of the world is originally configured as ethical experience, since our effort to assert ourselves as selves hinges on this fundamental relation to the other selves.

In way of conclusion, I shall try to argue for the primordial nature of such an ethical experience by looking on the configuration of experience as a practical space of reason rooted in an ontology of care.

Subjectivity in an Ontology of Care

In this concluding section, I shall try to gather the somewhat loose ends that the previous analyses have brought forth without drawing any general conclusion. Two notions will guide this attempt to round up the different analyses, namely the practical space of reason and an ontology of care[18].

The practical space of reason is the world experienced as a hierarchy of heterogeneous values generated by the self, the world, and the other. This space is configured as a unity of various practices that each have their own logic, but which are nevertheless still interrelated and organized according to the values that pertain to our existence as selves with an integrate identity through time. The narrative configuration of personal identity functions as a primary clarification of the practical space of reason because it explains the *how*, the *why* and the *who* of the experience of the world. It structures the otherwise confuse and heterogeneous character of experience by gathering together the various actions and sufferings in the temporal unity of a narrating self. Furthermore, it emphasizes the hermeneutic character of human existence. The values that configure our experience of the world are inscribed in a certain socio-cultural context that needs interpretation in order to become articulate and sometimes even conscious. It is a space of *reason* because it is always by means of the reflective capacity of reason that we articulate and cope with the alterity in and through which we seek to assert ourselves. Nonetheless, the narrative approach still embeds the subject firmly in the pathetic dimension of human existence. The subject is pervaded by alterity which is felt as 'the reduction, even the destruction, of the capacity for acting, of being-able-to-act, experienced as a violation of self-integrity' (OSA 190/223). The self-integrity of the subject is continuously challenged in the embedded and embodied existence in time. Reason works on the ground of this fragile integrity and must constantly seek to appropriate and assert the irreplaceable unity of the self by means of ongoing interpretations of alterity. Our experience is configured by this interpretation in the sense that we conceive of our existence a practical problem of how to cope with the alterity that challenges our identity. This results in practices that aim to enhance our wellbeing (i.e. self-assertion expressed in a search for identity) according to a hierarchy of values. A massive part of these values involves the presence and well-being of other selves in our own personal life-plans and our general idea of 'the good life'.

[18] Ricoeur's terminology is slightly unstable with regard to the notion of care. He uses the terms solicitude (sollicitude), care (souci), and the Heideggerian notion Sorge (the German word for care) interchangeably. I do not see any substantial difference in his use of the different terms, so I stick with the term care, unless one of the other terms feature in a direct quote.

Now, in way of conclusion, I shall argue that these interpersonal values have a privileged status in respect to other values and are therefore extremely critical for the organization of our hierarchy of values and, thus, for the configuration of subjective experience. These interpersonal values are all grounded in the phenomenon of responsibility which Ricoeur believes is not imposed externally onto the subject, but an intrinsic part of the structure of subjectivity. In the same move, I will try to show how the relation between selfhood and otherness, particularly visible in interpersonal relations, constitutes the ontological situation of the human subject.

The identity of the subject is, as we have seen, a complex problem that cannot be reduced to a question of numeric identity. Other than passively remaining the same (sameness/idem or character) through time, the subject must actively remain the same person (selfhood/ipse or keeping one's word). Thus, identity is also a normative question of integrity and self-esteem (OSA 167/195). Our identity as unique and irreplaceable selves, i.e., persons, depends on how our conduct (doing and saying) conforms to the values and norms shared together with other selves that are themselves unique and irreplaceable. The 'nonsubstitutibility' or non-fungible nature of the human self is a fundamental part of our personal identity (OSA 193/225). I am this unique person and not just a congregation of different qualities, actions, feelings, and dispositions. My existence, or more specifically my personal identity in this existence, matters, and much of my effort goes into asserting this personal identity through and in spite of time.

This normative feature of human subjectivity is fundamental for understanding the effort that Ricoeur makes to stress the particular ontology of human subjectivity. The subject is rooted in both physical nature and in a spiritual nature that he names humanity. The laws and rules that govern its existence are characterized by a polysemy that prevents a straightforward understanding of the meaning of human action and suffering, and 'which authorizes no more than the idea of an analogical unity of action' (OSA 312-3/362). Our actions (and sufferings) do not only follow the rules of physical nature or unconscious desires (archeology), but find a part of their constitutive meaning and motivation in the desire to be inscribed in a human society, that is, humanity (teleology) (OSA 346/399). The being of the human subject is characterized by this intersection of physical and spiritual laws, values, rules, and norms which creates the tension or fragility in the core of the subject and finds its expression in the vulnerability to conflicts that haunts the self-assertion of the subject. The meaningful world experienced by the subject is not exhausted by a physicalistic explanation, but requires a more open-structured approach that does not decide to hasty on the nature of human subjectivity but rely on an interplay between self and otherness: 'The

polysemy of selfhood, the first to have been remarked, is revealing with respect to the polysemy of the Other, which stands opposite the Same, in the sense of oneself' (OSA 318/368). The human self is in a particular ontological position because, on the one side, it is firmly rooted in physical nature, and on the other, it seeks to assert itself through, and often in spite of, that nature. Alterity and selfhood are interrelated in a continuous tension that results in an 'unusual ontological situation […] this situation is the object of an *attestation which is itself broken*, in the sense that the alterity joined to selfhood is attested to only in a wide range of dissimilar experiences, following a diversity of alterity' (idem. Translation slightly modified). We have seen that this alterity plays a constitutive role on every level of subjectivity, from the body over the world to the other subject. The ontological status of the human subject is constituted in this interplay between selfhood and alterity in the sense that the subject becomes a self through an interaction with the alterity in which it is rooted and situated. This interaction is shaped as a configuration of the subjective experience of the world, the other, and the self through language, society, history, institutions etc. Alterity becomes meaning, objects, things, and other persons through that configuration. The subject must find itself, assert itself, through the configuration of that which is other. This, however, is not the same as to say that the subjective world is what the subject makes it into. One must remember that selfhood is not simply mineness. The experience of the autonomous nature and value of otherness (the body, the world and the other) excludes the possibility of solipsism and emphasizes the pathetic dimension of subjective experience: we act *and* suffer, experience *and* feel, think *and* sense. To understand more clearly this 'unusual' ontology that Ricoeur aims at, we have to look, one last time, at the peculiar experience of alterity represented by the other self.

The most compelling experience of alterity is that of the other self. This is not to say that the experience of body and world is not compelling in its multifarious forms of manifestation, only that the other self plays a more dynamic and precarious role in the subject's self-assertion. The other responds to our action in a far more complex ways than the other manifestations of alterity. The other self is an active part of our identity as persons. The presence of the other is not something externally imposed on the existence of the self, but an intrinsic part of the constitution of selfhood. The subject experiences the alterity of the other self as a need or a lack in its normative idea of 'the good life': 'To find oneself called upon in the second person at the very core of the optative of living well […] is to recognize oneself as being enjoined to *live well with and for the others in just institutions and to esteem oneself as the bearer of this wish*' (OSA 352/406). The attestation of selfhood is intrinsically interlaced with the experience of injunction in the constitution of personal

identity. When we assert ourselves as selves we find our identity as persons dependent on the response of the other. We need the response of the other in order to maintain the integrity of identity. This need is experienced in self-assertion as the passive phenomenon of *conscience* (OSA 351/404-5). We cannot escape the presence of the other in our self-esteem and our wish for 'the good life'. The other is part of the experience of ourselves, since the legitimacy of our doings and saying is always weighted against the presence of the other in the attestation of our own self. In the experience of conscience, we experience 'being enjoined as a structure of selfhood' (OSA 354/409. Translation slightly modified). We simply care for the other's opinion, because that opinion is constitutive for the integrity of our personal identity. Faced with the other, we are held responsible for our sayings and doings in our existence over time (keeping one's word). The other self grounds our actions in common human world that is not constituted by my own initiative. My care for the wellbeing of the other self is experienced in the *need* of legitimacy that characterizes my motives and dispositions to interact with the world and the other. Conscience is where this need manifest itself. My identity as a person cannot be sustained in a constant disregard of the other subject. Although the voice of the other is not actually present, the otherness experienced in conscience makes this voice fundamental for the integrity of my identity, for my self-esteem. Ricoeur can therefore say that care gives ontological weight to the conception of human practices (OSA 311/361). The human self cannot be explain by an ontology that disregards this fundamental relation between care, conscience, and attestation, since the self is 'a concept that only an ontology of care enable us to constitute' (OSA 188/220). We care for ourselves, i.e., the attestation of ourselves as an individual self through time (personal identity). In this attestation, we are enjoined by the presence of the other to take care of the other in the appropriation of our own selfhood. This being-enjoined is experienced as a need for legitimacy of our actions, since we are responsible for those actions in front of the other self. The existence of the self is not conducted in isolation but as an integrate *person* among other persons. Our self-esteem (the reflective instance of 'the good life') is constituted by the basic care for the other and solicits a spontaneous desire to obtain the approval of the other, a recognition of my integrate identity as a person. It is on the basis of such an ontology of care that it is important for Ricoeur to emphasize that the question of responsibility is not something external to the structure of human subjectivity; on the contrary, it resides deeply in the constitution of subjectivity: 'This is why it is so important for us to give solicitude a more fundamental status than obedience to duty. Its status is that of *benevolent spontaneity*, intimately

related to self-esteem within the aim of the "good" life. On the basis of this benevolent spontaneity, receiving is on an equal footing with summons to responsibility' (OSA 190/222).

Personhood and responsibility is therefore intrinsically intertwined. The identity of the subject as a person is constituted and maintained by the responsibility for the other because it is also responsibility for itself as a person: 'Because someone is counting on me, I am *accountable for* my actions before another. The term "responsibility" unites both meanings: "counting on" and "being accountable for." It unites them, adding to them the idea of a *response* to the question "Where are you?" asked by another who needs me. The response is the following: "Here I am!" a response that is a statement of self-constancy' (OSA 165/195). This intrinsic relation between personhood and responsibility has a substantial impact on the ontological status of human subjectivity. The being of the human subject does not allow for a strict distinction between fact and value. Human nature cannot be envisaged and explained with taking into account the normative dimension of its being. Normativity is not something that is added on to the factual nature of the subject in a second move, but the very core of its being. The self is not static but always on the way to appropriate itself as an identical, irreplaceable, and integrate self, i.e., a person. Selfhood, as have been emphasized from the beginning, is marked by a non-coincidence that often is experienced as a conflict between the self as a part of physical nature (vital desires) and as a part of humanity (spiritual desires). This conflict is ultimately a conflict between selfhood and alterity that is expressed in the attestation of the identity of the self through the world, the body and the other self. The being of the self is made up by alterity and its identity as a person and cannot be understood in isolation from this otherness. The self is rooted in the alterity of physical nature (the body and the world) and in the alterity of the other selves (humanity). The values that make up the existence of the self are determined by these different experiences of otherness other than by its desire for self-assertion.

In chapter one, the structural analyses of the subject uncovered a tension in the self between an originating affirmation and an existential difference that solicits conflict in the heart of the subject between vital and spiritual desires. These analyses led to an outline for an ontology rooted in a continuous encounter between selfhood and alterity. The human subject is fragile and vulnerable to conflict due to this complex nature of its being. The subject is not at rest but in constant search of restoring the identity of itself as a self. This restoration of itself is experienced as a conflict between the selfhood and alterity, since personal identity is constituted not only by its own self but also by the alterity that continuously challenges (and threatens) this identity in the course of its existence. The feelings that qualify its existence and motivate its interaction with alterity (the body, the world,

and the other) are structurally differentiated in vital and spiritual desires. On the one hand, it relentlessly seeks to affirm or assert its (well-)being at every cost. On the other, it needs alterity, especially the other self, because its existence is an existence in and through that which is different from itself. Now, Ricoeur shifts to a hermeneutical method because the conflict in the heart of the subject cannot be solved at a structural level, since the felt alterity is crystallized through the subject's practical interaction in the world through time. The subject is lost in the embedded and situated nature of its existence in time. Alterity manifests itself through history, society, culture, institutions etc., and these different manifestations need ongoing interpretations to be appropriated as the world of the subject. This interpretation is always rooted in care or a concern for itself *and* for alterity. That is, the subject cares for the other, although it seeks to assert itself in and through that alterity. Care is what binds together the different aspects of the subject.

This ontological status of care is particularly clear with respect to the other self. The subject cares for its place among other selves (its idea of 'the good life') and feels a strong desire to be part of a human society. This desire for humanity is fundamental for its identity as a person among other persons, because it needs the other self to affirm its identity as an integrate and irreplaceable person. Self-esteem is part of our identity, and the other's recognition is critical to how we esteem ourselves as a person. We are responsible to the other self because the other self responds to our actions, and we, on our part, act and suffer in accordance with that response. We care about what the other have to say about our person. Therefore, our identity as persons depends on the recognition of the other self other than on the mere factual continuation of our being.

The particularly critical importance of the other self is the reason why ethical values are dominant in the configuration of subjective experience. The basic care that animates the subject's interaction with the world is particularly emphasized with regard to the other self, since the other self is a fundamental part of our personal identity. Our actions are characterized by an unrelenting claim for responsibility, which is rooted in the structure of our being because we need and care for the other's recognition. Ethical values and norms are therefore needed in order to assert our selves as persons in a world shared with other persons. Our coexistence with the other persons is often permeated with conflict, disagreement, and dilemmas, which solicit this need for an ethics. To assert itself, the subject must know how to interact with the other self. The question of normativity is not an extra dimension added on the factual dimension of subjectivity, but inscribed in the nature of being a self. An ontology of care attempts to embrace this complex structure of human subjectivity by emphasizing the relation between selfhood and otherness in the constitution of the subject.

Conclusion: A Framework

This attempt to reformulate Ricoeur's theory of subjectivity is a first, necessary, step in the investigation of the relation between subjectivity and ethics. I could have chosen to attack the problem directly without first consulting Ricoeur's work. I believe, however, that such an approach would start on a less substantial and clarified ground that I do now. Thus, I deem the foregoing reformulation necessary.

The following analyses will use the reformulation as a framework for the question about subjectivity and ethics. It will depart from Ricoeur's analyses in decisive ways. I use a lot of time of the concept of human affectivity, which in Ricoeur played a critical role in the stripped notion of subjectivity, but was somehow left out or only peripherally touched upon in his later works. Further, I articulate the neurophysiological dimension of emotions in order to understand the embodied (biological) nature of human feeling. Ricoeur did not do this, and normally he did not involve biology in his analyses. His dialogues with the neurobiologist Jean-Pierre Changeux (WT) shows that he does not exclude that we can learn something about the subject from neuroscientific research. However, he continually recommends that we do not draw hastily conclusions or perform immature reductions of subjective experience to subpersonal mechanisms of body and brain.

In fact, the following pages will be an extensive analysis of the relation between subjectivity and ethics in order to argue for a non-reductive naturalistic conception of human personhood. I extend the arguments emphasized in the reformulation, namely that subjectivity and ethics are irreducibly related, and that subjective experience is configured as a practical space of reason. I then claim that this is not contrary to a neuroscientific approach to human nature if such an approach respects the irreducible nature of human subjectivity. I ground this claim in the two notions that were brought forth in the foregoing reformulation, i.e., the practical space of reason and an ontology of care.

Part Two

Feeling Ethical

The subject cares about being a person, just as she or he cares about so many other things. However, personhood stands out among the things that we care about; maybe not explicitly in our daily life, but whenever our personhood is injured, threatened, or offended we feel how much we care about being the person that we are. Somehow, the person that we are characterizes what we care about, or how we care about the things that we care about. Personhood remains the firm, but obscure, point of reference when we talk about ourselves and other people. A person can be humiliated, loose his or hers dignity, meaning of life, respect for other persons, change character, become a better or worse person. People can be treated as if they were not persons (most of us in the western world learn this from history and television, but too many suffer this horrifying experience). We can kill, torture, humiliate, starve and laugh at a person, but we cannot kill or torture personhood out of a person. Persons, as Richard Wollheim writes, 'have something in common, which is membership of a particular biological species, and, since we also know no human beings who are incontestably not persons, it is easy to conclude that to be a person is to be a human being. They are one and the same thing' (1984: 3). This is an interesting feature about human subjectivity. We are persons by the simple fact that we breathe and exist as a member of the human species, and still we care about being persons. We care about being something that we already are.

The reformulation of Ricoeur's theory of subjectivity insisted on an ineradicable relation between subjectivity and ethics by emphasizing the critical presence of the other in the experience of selfhood. The subject's experience, action, and feeling are constituted by the presence of other subjects to such an extent that subjective experience is not constituted primarily by the existence of a self, but by a coexistence together with other selves; that is, the existence of the self as a *person* embedded in physical nature and situated in a human society inhabited by other persons. I therefore introduced the notion of a practical space of reason to argue for a configuration of subjective experience as a hierarchy of heterogeneous values generated and refined by the self, the world, and the other. Further, I briefly touched upon Ricoeur's idea of an ontology of care as the fundament for such a configuration of experience. The being of the human subject is characterized by what that particular subject cares about.

This part is dedicated a further exploration of the problematic relation between subjectivity and ethics. The analyses, however, rarely involve ethics as an explicit theme, and, when I finally do this in the second chapter, is might be characterized more as metaethics than ethics. This is done intentionally, since my main interest in this work is to clarify why and how subjectivity and ethics are related and not ethics in itself. I shall argue that subjective experience, feelings, values, and practices are primarily ethical and conclude that we cannot not understand properly neither subjectivity nor ethics without understanding how and why this is so. Personhood is essential to this conclusion. This will be alluded to throughout the analyses, but first clarified more in detail in the final sections of the second chapter.

The analyses will be carried out in two main sections. The first chapter looks on the affective dimension of subjectivity. Human affectivity marks the interface between the physiological and the intentional aspect of human subjectivity and will therefore require a treatment of both these aspects. The second chapter deals with ethical experience. I analyze the normative aspect of subjectivity and argue for the constitution of human experience as a practical space of reason rooted in an ontology of care.

Chapter one

Human Affectivity

Human affectivity is mainly experienced and expressed as emotions. We are moved by the way that the world affects us which has therefore made emotions the central theme in the study of human affectivity. Nonetheless, emotions have had their ups and downs for little more than a century now. In the end of the 19th century two important thinkers occupied themselves extensively with the nature of emotions, namely Charles Darwin and William James, but for the better part of the 20th century emotions have lived a quiet existence in philosophy. Most philosophers and psychologists agreed that emotions are fundamental for human existence, but they nevertheless chose not to deal with them explicitly. However, in recent years emotions have once again come into focus in philosophy as well as in other branches of cognitive science. This might be a consequence of the growing interest dedicated to the phenomenon in the empirical sciences, or maybe just one of the many turns inside philosophy itself. Nonetheless, today at least one thing seems certain: a philosophical investigation of emotions cannot be done seriously without bringing into account the empirical investigations of the biological nature of human affectivity.

In this section, I will first look at the conceptual background for the philosophical study of emotions. I present three main theories: *the feeling theory*, *the cognitive theory,* and *the narrative theory*, since these give a pretty good idea of the contemporary debate. I briefly sketch the general features of the three theories so that I can prepare the ground for my own account. However, before turning to my own considerations, I shall look at the neurophysiological dimension of emotions that I retain a fundamental element in any attempt to approach human emotions. Then I finally argue for an alternative account of emotions that is inspired by Ricoeur, nourished by neurophysiological insights, and directed by the groundwork done by the three theories in contemporary debate. I claim that these theories neglect the importance of moods in emotional experience and argue for a theory that accounts for the neurophysiological dimension of emotions and furthermore differentiates feelings, moods, and affects as different aspects of the same experiential phenomenon.

Conceptual Background

The Feeling Theory

This theory is perhaps the most representative of common sense, namely that emotions are special kind of felt experiences like sensations (hearing and seeing) and proprioception (registration of

changes in the body). However, emotions somehow differ from other felt experiences by means of the qualitative feature of their manifestation; they are vested with a certain felt quality.

The theory found its classical expression in the James-Lange theory of emotion. The two scholars, independently, proposed a theory that reduced emotions to perceptions of bioregulatory changes in the body. James' famous example of fear of the bear states that common sense tells us that we see the bear, become afraid, and then run. James turns this picture up side down saying that this is all wrong. We fear the bear because we try to run from it. Emotion is not a result of our cognitive relation to the bear as if I see a bear, know that it is dangerous, become afraid and run. On the contrary, our emotions are the experience of feelings caused by the perception of a set of bodily responses. I see the bear, my body reacts to the danger that I perceive, and I have the feeling of being afraid because I perceive the visceral reactions of my body. James writes: 'What kind of an emotion of fear would be left, if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think' (James 1884:193-4).

During most of the 20th century the theory has been criticized for not being able to distinguish between different kinds of emotions caused by the same visceral reactions. The first important criticism, carried out on a physiological basis, was expressed by W. B. Cannon who in 1927 argued that 'the responses in the viscera seem too uniform to offer a satisfactory means of distinguishing emotions which are very different in subjective quality. Furthermore, if the emotions were due to afferent impulses from the viscera, we would expect not only that fear and rage would feel alike but that chilliness, hypoglycemia, asphyxia, and fever like them' (Cannon 1927: 110). The same bodily state may underlie two different emotional experiences such as fear and rage. The phenomenology of fear and rage differ in such a radical manner that the difference cannot be accounted for by visceral changes alone. There is obviously something more to the emotions than felt bodily changes. After this frontal attack, the feeling theory seemed surpassed in the philosophical and psychological study of the emotions.

Nevertheless, about a decade ago the theory gained new force with the compelling studies of the neurobiologist Antonio Damasio (1994, 1999, 2003). Whereas James and Lange came to their idea of emotions as feelings mainly from the visible physiognomic expression of emotions, Damasio scrutinizes emotions inside the brain, that is, in the neurophysiological properties and dynamics of emotions. He concords broadly with the James-Lange theory (Damasio 1994: 129) and develops their thesis by reflecting on new evidence provided by a combination of clinical observation and

cognitive neuroscience. He convincingly defends the feeling theory from the criticism launched by Cannon. Among other things, he shows how the body is not always distinctly present in the feeling of an emotion. The brain is able to somehow 'by-pass' the body by means of what he calls 'as-if'-feelings: 'we conjure up some semblance of a feeling within the brain alone.' (Damasio 1994: 156). And yet he immediately emphasizes that 'I doubt, however that those feelings feel the same as the feelings freshly minted in a real body state' (idem). He surmises that there is not always a direct line between a specific body state and a certain feeling of emotion. He differentiates the various feelings and works out an elaborate account of the idea that emotions are feelings, going from simple homeostatic regulations to acquired, social emotions. For example, he gives a more detailed account of the feeling of sadness than the one we find in James who just emphasizes the physiological machinery of sorrow: 'each fit of sobbing makes the sorrow acute, and calls forth another fit still stronger' (James 1884: 197). Damasio claims that alongside the perception of the body state there is the perception of certain thoughts that might just contribute to the feeling of sadness (Damasio 2003: 85). These thoughts are metarepresentations of our own mental process that allow the mind to represent other parts of the mind (metal images and long-term memory). In this way, Damasio tries to argue for the variety of emotional experience by emphasizing the complexity of our perception: '*a feeling is the perception of a certain state of the body along with the perception of a certain mode of thinking and of thoughts with certain themes*' (Damasio 2003: 86).

In order to develop this detailed version of the feeling theory, he has investigated further into the subjectivity of emotional experience, but never ceasing to emphasize the fundamental importance of the homeostasis, i.e., the ebbs and floods of 'the bodily landscape', his notion of subjectivity remains somewhat simplistic. And this simplicity affects his theory of emotion. A subject is what it is, and nothing else. It is a core self that, throughout our existence, develops into an autobiographical self, shaped by education, culture, and unique and contingent personal events. Theses factors create, in a sort of orchestration, the person that we are. Naturally, he does not exclude our capacity to create imaginative variations on our actual being, but he seems certain that, in the end, we are the entities that we are constituted to be: 'We can be Hamlet for a week, or Falstaff for an evening, but we tend to return to home base' (1999: 225). This idea of a home base underestimates the reflexive aspect of subjectivity, i.e., our capacity of to evaluate and critically revise our actually being the person that we are.

I return to Damasio's account again in the next section, but first we shall look two competing theories that criticize the reduction of emotions to pure bodily feelings. Emotions play an important

role in our evaluations and cannot be reduced to the feelings of bioregulatory changes in the body state. Our emotions have an intentional nature that surpasses the internal feel of body states. It is this intentional nature that the cognitive theories have leveled at in their criticism of the feeling theories.

The Cognitive Theory

Most philosophical theories of emotion are a in some way or another cognitive theories. They claim that emotions involve propositional attitudes. When I am angry, I am normally angry at someone because of something. When I am happy or sad it is normally due to something that I am aware about. Emotions have a logic (Solomon 1976) or a structure (Gordon 1987) that surpasses the perception of bodily changes. They are judgements about the world and my situation in it, involving a more global stance towards my existence. Feelings may occur together with emotions, but as one of the most enduring defenders of the cognitive theory, Robert C. Solomon, writes: 'one can have an emotion without feeling anything, and one can feel anything (including all of the 'symptoms' of emotionality, for example, flushing, pulsing) without having any emotion whatever' (Solomon 1976: 161)[19]. What Solomon emphasizes is that emotions reveal our intimate relation to the world and to ourselves, not only as possessors of a complex primate brain and an autonomic nervous system, but as persons with certain ideas, beliefs, and commitments not reducible to the internal feel of body states: 'every emotion establishes a framework within which we commit ourselves – or refuse to commit ourselves – to our world and to other people' (Solomon 1976: 202). Most cognitive theories consider emotions as judgements that have a specific logic or rationality (Sartre 1971; Nussbaum 2001: 19-85; Taylor 1985: 60-2). This logic or rationality is sometimes identified with the complex concept of intentionality in the sense that emotions are expressions of certain beliefs or desires. When I feel irritated in a given situation it is because of something going on in this situation. I am irritated at the person because of something the person said or did. Perhaps his statement makes my opinion look ridiculous or perhaps just because he moves himself in a stupid way. I am jealous because my neighbour has an expensive car or his apple tree carries more delicious apples than mine. The combinations are endless, but it is a question of structure. There is an object or proposition of the emotion. In order to understand the different emotions, we must try

[19] Solomon is mainly inspired by Nietzsche, Sartre, and the existentialist tradition. He wrote his first work as a polemic attack on the feeling theory (1976). But in the last years up to his death in 2007, he came to modify his theory in the sense that he made it less polemical and included feelings as an important part of emotional experience. Nonetheless, he still emphasized the primary importance of the cognitive, evaluative stance (2001: 189-192; 2007: 205-6).

to sort out the different structures or complexes of this object relation. This maybe followed by feelings, but the feelings are secondary. What is primary is the propositional attitude, belief or desire expressed in the emotion. The emotion is explained and defined by scrutinizing the propositional attitude what ever this may imply: beliefs, judgements, desires or complexes of these.

Feelings are in this scenario reduced to 'add-ons' (to use an expression from Peter Goldie, 2000a: 40-1). They are not part of the explanation of the logic of emotions. This may sound familiar. Actually, feelings seem to suffer from the same over-intellectualization that for a long time excluded emotions from philosophical investigations. Russell wrote in 1927: 'The emotions are what makes life interesting, and what makes us feel it important. From this point of view, they are the most valuable element in human existence. But when, as in philosophy, we are trying to understand the world, they appear rather as a hindrance' (1960: 228). It is the same distinction that is at play in the cognitive theory of emotion. Feelings are what makes emotions interesting and what makes us feel them important, but they are a hindrance when we try to understand them.

It is glaringly obvious that such theories have been met with much skepticism (Goldie 2000a: 74-78; de Sousa 1987: 40-1, 165-6; Stocker 1983: 18-23). How can one explain emotions without including the feeling aspect? One of the classic arguments against the cognitive approach is Michael Stocker's 'fear of flying' objection (Stocker 1996: 38-9). We know that flying is the safest way of transportation and still we may be afraid to fly. This may be explained as an irrational emotion, but cannot be solved by showing that statistics demonstrate the belief to be untrue. Changing the belief will not change the emotion. And we do not consider a person an idiot for being afraid. It is just how airplanes make her feel. It may be an inappropriate feeling but it may still cause her to choose others and more tedious forms of transportation in order to avoid traveling in an airplane. Feelings cannot be rejected as part of the emotion because of the simple fact that the non-logical feeling often causes the person to act in a certain way.

The Narrative Theory

In recent years, a third theory is slowly coming in to shape. Or more correctly, an old view on emotions has taken a new form. It is an attempt to combine feeling theory with some aspects of the cognitive theory because '[f]eelings are, as we all know, at the heart of emotion' (Goldie 2000a:12), and yet emotions, and the feelings involved, have a certain structure of intentionality. Emotions are not reducible to beliefs and desires, and still they caused by more than mere internal body states. This position has been prepared by the late Richard Wollheim (1987, 1999) and is under continuous

elaboration by Peter Goldie (2000a, 2002ab, 2003abc, 2004, 2007). The general idea, however, is not new. It draws on insights from Freud, Heidegger, Sartre, Ricoeur, Charles Taylor, and Solomon. It emphasizes the personal aspect of emotional experience. An emotion remains unintelligible if we try to approach it from an impersonal stance (Goldie 2002b: 249). Emotions tell us something about how the person experiences the world, itself and other people. Where a scientifically inspired philosophy aims at explaining the impersonal aspects of the human subject and the world, a philosophy of emotions, unavoidably, has to deal with how the person experiences the 'magical world' of emotional experience. Sartre describes emotional consciousness in the following way: '[it] is a new consciousness in front of a new world – a world which it constitutes with its own most intimate quality, with that presence to itself, utterly non-distance, of its point of view upon the world' (Sartre 1971: 78/42. Translation modified).

In the new approach, however, this personal stance becomes the centre of a narrative structure that holds together the different elements of an emotion: past episodes of emotional experience, imagination, desire, belief, character traits, and dispositions to feel, think, and act in certain ways. It reacts against the belief-desire account of emotion arguing that whereas a belief or a desire can be feelingless, an emotion always involves feelings as fundamental part of its nature. For example, I may have the belief that the fjords in Norway are freezing cold, but do not feel it unless I go to Norway and dip my toe in one of them; I may desire to go on a tour around the world without experiencing it as an emotion. This is not to say that beliefs and desires do not carry any important relation to emotion. The theory simply emphasizes that emotions cannot be reduced to such accounts. The nature of an emotion is a complex affaire, which is best approached as a narrative structure, or in Goldie's words: 'The best understanding of a person's emotions will take account of the distinctions I will be discussing [Bodily feelings, feelings towards, dispositions, mood, character traits, and so on], but will nevertheless be holistic in its overall approach, seeing feelings as embedded in an emotion's narrative, as a part of the person's life' (2000a: 51). This way of investigating the emotions (which I have named the narrative theory) consider emotions in how they are experienced by the person, which is – phenomenologically speaking – a complex phenomenon constituted by intentional feelings, dispositions to act, and moods. These complex experiences are then articulated by language and the stories in which they are embedded. As Charles Taylor writes: 'To experience an emotion is to be in a sense struck or moved by our situation being of a certain nature. Hence, I said, we can describe our emotions by describing our situation' (Taylor 1981: 107). The most intriguing feature of the narrative theory is that by understanding emotions to be of an

essential narrative nature we obtain a tool to investigate more accurately into the ancient relation between emotions and ethics (a thesis developed further in chapter two).

One critical problem about this theory is that its defenders do not account for the biological dimension of emotions. They are too preoccupied about defending the intentional and personal aspects of the emotions that they neglect the fact that emotions are the most profoundly embodied phenomena of human experiences (de Sousa 1987: 47)[20].

The three theories that I have outlined above emphasize three important concepts that will serve as guidelines in my own account of human affectivity: 1) feelings, 2) cognition (intentionality), and 3) narratives.

But before I turn to this account, we have to look more closely on the embodied nature of emotions. As I mentioned earlier, today a theory of human emotion must take into account, at least to a certain degree, the biology of emotions. The following section will present two neurophysiological theory of emotions. I begin with a brief presentation of Damasio's work on emotions and feelings, since his studies of emotional consciousness have much of the credit for introducing neuroscientific research into the philosophical debate of emotions (some might even say that it was his work that revived emotions in philosophy). However, it is the work of another neuroscientist that I believe contributes most to the philosophical study of emotions, namely Jaak Panksepp's pioneer research on the affective dimension of the mammalian brain.

Neurophysiological Dimension of Emotions[21]

The complexity and importance of human emotions are deeply involved in its embodied nature. The phenomenology of emotions cannot be severed from the ontogenesis of these ubiquitous phenomena of human life[22]. The biological nature of emotions is too compelling to be passed over

[20] In fact, Goldie argues for 'a profound distinction between the science of human nature and the natural sciences' (2002a: 98) in at least three respects. Human nature is characterized by perceptive, normativity, and historical condtion (106-7), which cannot be explained by the natural sciences, since they adopt an impersonal perspective (cf. Goldie 2002b: 249-50; Goldie 2000a: 181-2). I largely agree to this, but I will still insist that the distinction is less profound than Goldie makes it.

[21] I have found it necessary to avoid specific neuroanatomical and neurophysiological terms to the extend that I deem defensible, since the general ideas of the theories may be described without the technical detail. The anatomy, physiology, and chemistry of the brain are complex matters and require, to be only minimally understood, a long preparatory introduction that is not possible here. However, the technical detail is important, and in order to fully understand the scope and argument of the theories one must turn to the references indicated in this section. A fine introduction to the topic is found in (Bear, Connors and Paradiso 2006).

[22] Ricoeur emphasizes that it is the affective dimension of subjectivity that evidences the ontological dimension of subjective experience (FM 103/119; FN 88/84; cf. pp. 45-50 above). The subject's being in the world is revealed by the feelings; it *is* already in the world before it comes to *know* about its being in the world. The particular being of

in silence. The fact is that my body is often the first indicator that I am experiencing a certain emotion or find myself involved in an emotional episode. I feel a tension in my stomach and uneasiness in my limbs during a conversation with a certain person; I may be drumming my fingers on the desk or moving my feet rapidly under the table. And then I come to realize that I am intensely irritated by or with something about that person. Bodily feelings are often involved in the expression and experience of emotions.

One way to investigate the role of the bodily feelings in emotions is to look at the neurophysiological properties and dynamics involved in emotional experience. In the last decade or so, three scientists in particular have set the agenda for the bridging of philosophy and the neuroscientific study of emotions: Antonio Damasio, Joseph LeDoux and Jaak Panksepp. They have all three done extensive research on various aspects of the neuronal underpinnings of emotions. Damasio's most important contribution concerns the critical role of emotions in human decision-making (1994) and the relation between consciousness and emotion (1999), whereas LeDoux has worked mainly on the basic emotional mechanisms that are common to both humans and other animals (in particular insights on emotional learning and emotional memory derived from the study of rats) without including the subjective dimension of emotions (1996). Panksepp, on the contrary, is highly concerned about the subjective dimension (the feeling) of emotions in the mammalian brain. He differs from both Damasio and LeDoux (who both mainly deal with the cortical elaboration of emotions) by stressing the subneocortical dynamics of affectivity (1998). In the following, I will briefly present Damasio's and Panksepp's theories of the neurophysiological dimension of emotions. I argue that Panksepp's research is more congenial with a philosophical approach to human affectivity because of its emphasis on the subjective experience of emotions. But although I disagree with Damasio's general methodology, his research on the importance of feelings in decision-making (the so-called 'somatic-marker hypothesis') remains decisive for the analyses of ethical experience in the next chapter. Ledoux's research deserves to be accounted for, but for the focus of the analyses I have chosen not to do so.

Antonio Damasio: Objective Emotions and Subjective Feelings

Damasio's contribution to the philosophical investigation of emotions remains indispensable. He has convincingly shown that the feeling aspect of emotion is fundamental not only in our

---

subjectivity is determined by our being a subject inexorably rooted in the biological nature of our body as an object among objects.

understanding of emotions but in all cognitive processes. He concentrates on what he has named the somatic-marker hypothesis. By and large, the somatic marker consists of gut feelings 'that increase the accuracy and efficiency of the decision process' (1994: 173). Examining neuropathological cases Damasio has come to the conclusion that "cool" reasoning is not enough to trigger a decision. A person would never be able to decide on anything without the help of bodily induced emotions. He would be in the undesirable state of infinite possibilities of equal (or no) importance. The bodily induced emotions are what highlight some possibilities and not others. The emotions are thus what qualify our experience as a set of affordances that present themselves to us as feelings of what is the most convenient, appropriate, or reasonable among different possible choices.

He has come to this insight from the study of, among others, a young man in his thirties that he, for the sake of discretion, has chosen to call Eliot. Eliot was intelligent man, a good husband and father, who tragically developed a large tumor that ended up covering both frontal lobes. Fortunately, the tumor was benign and successfully removed by surgical intervention. However, during the intervention much of the frontal tissue that had been damaged was completely removed. This had a critical impact on Eliot's life. He had undergone a radical change of personality, even though his mental skills and memory seemed completely intact. He became unrecognizable in the eyes of his friends and nearest family: 'Eliot was no longer Eliot' (idem: 36). He was never happy nor sad, incapable of decision and unable to focus on longer-term projects. Furthermore, he always remained uncommonly controlled and indifferent even when he told the tragedies of his life (divorces, unemployment and bankruptcy). As Damasio writes: 'I began to think that the cold-bloodedness of Eliot's reasoning prevented him from assigning different values to different options, and made his decision-making landscape hopelessly flat' (idem: 51). One might say that reason alone can never make a choice. Experience has to be somatically marked in terms of feelings in order to guide stable conscious choices. In this sense, feelings are what makes our life matter, and the qualification of our experience resides precisely in this capacity of feeling the world and ourselves: 'consciousness *feels* like a feeling, and if it feels like a feeling, it may very well be a feeling' (1999: 312).

On the basis of the somatic-marker hypothesis (i.e., the importance of feelings), Damasio has over the years developed a fine-grained neurobiological theory of the emotions. On the following pages, I will present a very rough outline of his main ideas on emotions and feelings, and furthermore argue for the reasons why I choose Pankseep's model to his.

Damasio's account is based on a somatic theory of emotions. As mentioned above, it is a variant of the James-Lange theory that emotions are basically feelings of bodily changes. There are, however, two important contrasts to the earlier theory. The first is the already mentioned possibility that feelings can be activated endogenously by 'as-if' loops that bypass the body and simply stimulate somatic brain areas (idem: 156). In this way, we can have feelings without actual bodily reactions, which saves both time and energy (1999: 281) and procures a more nuanced picture of the feelings (not all are directly involved with the body). The second contrast is his emphasis on unconscious emotions. The brain can register alterations in the bodily landscape without this resulting in conscious feelings: 'emotions can be induced in nonconscious manner and thus appear to the conscious self as seemingly unmotivated' (idem: 48). For example, we may smile or sob spontaneously without knowing why (1994: 140-2; 2003: 70-3). But, by and large, Damasio sticks to the general idea of emotions as constituted primarily of bodily changes (e.g. visceral, musculoskeletal, and hormone alterations). The body is a 'theatre of the emotions' (1994: 155), because our emotional responses always originate, in some way or the other, in the 'internal mileu' of somatic reactions. These physiological responses, triggered by certain brain systems when the organism represents objects and situations, are part of the bioregulatory devices fundamental for the survival and well-being of the organism and can basically be reduced to regulations of the automatic 'homeostasis machine' that approaches or avoids certain objects (2003: 30-31) in accordance with pleasure and pain (2004: 51).

His model is worked out by means of the methodological distinction between emotions and feelings. Although emotions and feelings are part of the same process, they are nevertheless two separate mechanisms (2000: 15). Whereas the emotions are specific and coherent collections of physiological responses, feelings are 'the mental states that arise from neural representation of the collections of responses that constitute an emotion within the brain structures appropriate for such a representation' (idem: 20). On the one side, we have emotions: consistent physiological reactions that span from the basic process of metabolism (automatic mechanisms of internal chemical balances) over simple pain and pleasure behaviors (reactions of approach or withdrawal) and drives and motivation (appetites such as hunger, thirst, or curiosity) to emotions-proper (e.g. happiness, sadness, disgust, surprise). On the other, we have feelings that functions as a mere 'read-out' of these physiological reactions, that is, as mental patterns that map or perceive bodily changes. Sometimes this mapping is a straightforward read-out; at other times it functions by means of the 'as-if-body-loop' that procures a pure neural map of the body state, but nevertheless it always

functions in reference to the body. Alongside the perception of the bodily changes, feelings do also depend on the subjective state of the experiencing organism: 'A feeling about a particular object is based on the subjectivity of the perception of the object, the perception of the body state it engenders, and the perception of modified style and efficiency of the thought process as all of the above happens' (1994: 148).

This leaves us with a picture of two separate physiological processes at the bottom of what we know as emotional experience, emotions and feelings. But whereas emotions are objective and public, feelings are subjective and hidden. Hence, we are able to observe and investigate the emotion, since these are engendered by measurable and clear biological changes, but the workings of the feelings remain in 'complete privacy' (2000: 15). We have to derive our understanding of feelings from what we are able to discover about the emotions. Damasio continues to produce elaborate accounts on how human beings come around to have such an enormous variety of feelings that they actually have, but every time he returns to the idea that feelings are always strictly bound to the body (even if the body is only virtually present) simply because 'the body is the main stage for emotions' (1999: 287). Therefore, he boils his treatment of the variety of feelings down to six universal feelings that are coined on the six universal emotions derived from facial expression and recognizability, namely fear, anger, sadness, disgust, surprise, and happiness (idem: 285)[23].

The criticism that Damasio now faces is the old question that Cannon leveled at James in 1927: can such a theory of feelings as perception or read-out of bodily changes account for the subtle and multifarious nature of actual human feelings? I believe not, and this is mainly because of the methodological distinction of feelings and emotions. If we operate with such a distinction (private vs. public, subjective vs. objective), we ironically end up with the same Cartesian distinction that Damasio fights so hard against (the title of his book from 1994 was *Descartes' Error*). Whereas Descartes distinguishes strictly between deanimated bodies (res extensa) and disembodied spirits (res cogitans), Damasio transfers this distinction inside the human being itself by talking about feelings as a mental read-out of a mindless body. The body becomes an entirely depersonalized object rooted so firmly in the theoretical soil of evolution that the conscious person's active, practical and thoroughly personal relation to his or her body is only seen as feeble and useless scratching on the harder surface of an objective and anonymous nature. In many matters, Damasio

[23] Damasio here conjoins the distinctions worked out by the famous research of the psychologist Paul Ekman who individuated these emotions as being universal across cultures and gender. Ekman defined the basic emotions according to several characteristics: distinctive universal signals, presence in other primates, quick onset, brief duration, automatic appraisal etc. (Ekman 1992: 175).

rightly observes, the body will have its ways notwithstanding the continuous struggle of the person. We have to understand the life of the body in order to understand the life of the person who has that body. Damasio warns us (as most paleontologists, neuroscientists, and evolutionary theorists) that the mind is only a conglomerate of newly developed brain systems deeply embedded in millions and millions of years of evolution (1994: 254). Therefore, we must understand that the body has its reasons, its own rationality (the working of the emotions) that is ultimately related to the need of survival (1994: 262-4).

Now, I will not in any way dispute the importance of emotions as embodied appraisals carried out in terms of registration of physiological responses in the body[24]; on the contrary, I do believe that emotions are intrinsically linked with the body, but it is the conception of the body and its presumed relation to the cognitive faculties that I resist. Ricoeur's analyses of the body individuated an ambiguity in our conception of the body (above pp. 70-8). The body is, at the same time, both *my* body and *a* body; that is, *my* personal body by which *I* access the world, i.e., a body characterized by my personality (when and where I was born, how I was brought up and educated, physical constitution etc.) and then *a* body as an autonomous object which is part of an anonymous nature that determines many of the reactions of that body, i.e., basically just a body among other bodies. This ambiguity of the body is missed in Damasio's (and even more Prinz's) account of emotions, and the cognitive element (judgment and intentionality) is therefore reduced to a somewhat strange, artificial elaboration of the real, basic emotions. The body has its reasons, but we cannot strictly separate these reasons form the nature of the person who feels and acts through that body. In the course of a life, the body becomes a part of the subject in that the subject is shaped by the body, but the body is also shaped by the subjectivity that it expresses. For example, on a pure neurophysiological level, the body of a skilled athlete differs only minimally from that of an

[24] Damasio's neurophysiological account of emotions has recently been defended and elaborated philosophically by J. Prinz (2004) who argues for emotions as embodied appraisals (77-8). All emotions are qualified by a kind of gut reactions that 'use our bodies to tell us how we are faring in the world' (69). Emotions that feel different from the basic body-related emotions are simply 'blends of basic emotions or cognitive elaborations of basic emotions' (93). For example, romantic love is a blend of the more fundamental emotions of attachment and lust, because 'some cases of love may be nothing than lascivious attachment' (124). This purely physiological approach to the emotions is then read back into the phenomenology of feelings. Prinz writes, for example, that 'I think the phenomenology of guilt is often just like the phenomenology of sadness […] The point is that certain emotions have similar or identical phenomenology. There are many other plausible examples: indignation feels like anger, disappointment feels like sadness, awe has an element of surprise, contempt has an element of disgust, pride feels like a kind of joy, exhilaration feels like a blend of fear and joy, and jealousy feels like a blend of anger, disgust, and fear' (2005: 19). All emotions have 'core relational themes' that are relevant to our needs and interests, in short our well-being (2004: 66), which means that fear, for example, is always somehow rooted in physical danger (219). This view on emotions as embodied appraisals determined by core relational themes have critical results for the role that emotions play in ethics. I return to this problem in the concluding pages of part three where I briefly discuss Prinz's ethical theory.

immobile academic, and yet the two persons have completely different experiences and expressions of the body. Their bodies are, at least in this sense, radically different from one another. A neurophysiological investigation of the emotions that operates with such a methodological distinction between emotions (body) and feelings (mind) will not be able to interact with a more phenomenologically based account. I mentioned earlier that Damasio's account of the emotions suffers from the simplicity of his view on human subjectivity, and this becomes evident in his reduction of feelings to perceptions of anonymous physiological reactions.

I have therefore chosen to use another neurophysiological theory of emotions as the background for my own account. Strangely enough, since I emphasized the personal aspect of emotions, this theory comes from a scientist who has spent more time with animals than with humans.

Jaak Panksepp: Subjective Affectivity

Over the last thirty years, Panksepp has argued for the need of more thorough animal studies if we are to understand the neurophysiology of human emotional experience: 'Through basic cross-species neuroscience research, the best we can presently achieve is the identification of essential neural components shared by all mammals, some of which may suffice to generate affective experiences – ancient experiential capacities I call *affective consciousness*' (2004a: 48). He insists that emotional experience is not generated only at the cortical level by newly gained neocortical cognitive capacities such as memory, attention, and cognition. He persistently levels a critique at the cognitive and behavioral neurosciences for ignoring the affective nature of human consciousness (2001: 136, 2003: 11-2, 2004b: 179; 2005b: 63-4; 2005c: 166-8). The picture of the human brain presented by the cognitive and behavioral neurosciences is envisaged on the ground of a neodualism that is even stronger than the old Cartesian one, because at least Descartes conceded animals to have various passions (1998: 340). This neodualism (present in both Ledoux and Damasio) believes that emotional experience (as well as human consciousness in general) is born only with the development of higher brain functions (the specifically human neocortex). Other animals do not have conscious experience and can therefore not feel emotions, but only react with instinctual emotional behavior (enabled by a combination of working memory and specific emotional systems concentrated in limbic system) (2007: 232). Therefore, this view holds that emotional experience is particularly human and must be sought out in the cognitive elaboration of bodily reactions. Accepting this presupposition, we end up with Damasio's sharp distinction between objective emotions and subjective (human) feelings as the only viable approach to the

nature of human emotion (2005a: 25). On the one side we have the ancient emotions developed and refined throughout the eons of evolution, and on the other the cognitive skills as newly evolved capacities (language, symbols, longtime memory, etc.) that in human have given rise to much more complex and elaborate emotional experience (feelings) and the consequent sophisticated behavior[25].

As a reaction to this traditional view, some twenty years ago Panksepp coined a discipline that he called affective neuroscience (1991). Contrary to the other kinds of neuroscience, affective neuroscience aims at explaining 'the existential reality of our deepest moods and emotions' (1998:14). The central working-hypothesis is that emotional experience, including subjectively experienced feelings, does play an important role in the chain of causal reactions that determine the actions of both humans and other animals. And, further, the subjective nature of feelings does not exclude scientific investigation. We must simply start by investigating the neurophysiology of the various subneocortical emotional systems that are homologous in all mammalian brains (2005c: 160). These systems generate different 'raw affects' or 'raw-feels' that instantiate intrinsic emotional values to orient the animal in the environment. Human and other mammal behavior originates from this interplay of different basic affective values. The functioning of the mammalian brain relies on a common affective consciousness, and cross-species research will help clarify the neurobiological underpinning of many basic human traits (2006: 781-2). The mammalian brain is basically affective (it interacts with the environment on the basis of certain felt values), and the higher brain functions that characterize humans cannot be severed from these basic affective value systems or retained to operate from 'top-down' read-outs of mere bodily affects (2004b: 179). The higher brain functions are themselves immersed in the affective dimension of ecological life, and to understand what is particular about human feelings we have to resort to a 'bottom-up' perspective that 'help clarify pre-positional affective psychodynamics that emerge from sub-neocortical regions of the brain' (2005c: 173). Combining the experimental advances of animal study (we can study other animals more thoroughly than human due to bioethical restrictions; e.g., it is not allowed to cut open a living human brain in order to take a precise sample of the sub-regions of the brain) and a phenomenological approach, we beget a new picture of the emotional brain, where 'emotional

[25] Panksepp, however, admires and largely agrees with the work of both Damasio and LeDoux (1998: 340; 2003: 7; 2004b: 181). What he resists is the methodological distinction between objective and subjective aspects of emotion. Whereas Damasio emphasizes the importance of feelings (although he individuates the higher cognitive brain systems to be the origin and thus maintains the dualism), LeDoux is even more radical and considers the conscious affective aspect of emotions (feelings) to be largely epiphenomenal. He writes, for example: 'Emotions evolved not as conscious feelings, linguistically differentiated or otherwise, but as brain states and bodily responses. The brain states and bodily responses are the fundamental facts of emotion, and the conscious feelings are the frills that have added icing to the emotional cake' (Ledoux 1996: 302).

affects may be thoroughly embedded within the extended activities of brain operating systems' (2005b: 38).

Panksepp calls this research strategy *dual-aspect monism*, which he considers the only way to challenge the classical behaviorism that is still very much alive in neuroscience (2005a: 24). We have to consider the complexities of the brain, causal and psychological properties, as being of one nature that nonetheless have to be studied from different (at least two) perspectives. The subneocortical affect systems that cause instinctual behavior and the higher cognitive functions that by experience and education develop and refine these instincts are two sides of the same coin and must be linked together in order to individuate 'basic dynamic features of brain and mind' (2005c: 163). This enables a scientific research of the causal foundation of affective consciousness in all mammals and not only the neural correlates of human cognitive elaboration of bodily emotions (idem). If we consider cross-species basic affects systems to influence and be present, with a constant variegating intensity, in all human cognitions and actions, we have a paradigm to study the neuroevolutionary mechanisms involved in the feelings that influence the being and action of humans. Neuroscience can thereby stick to observable and causal dynamics and properties of affective consciousness without jumping to speculations about the far more complex function of the neocortex. We cannot say anything even slightly definitive about the higher brain functions, let alone the way the different basic neocortical systems (e.g. auditory, vestibular, somatic sensory, and visual systems) conjoin to what we know as human conscious experience. So, although Panksepp's thesis about core affective consciousness in animals may seem an invalidate speculation since we cannot verify such a seemly radical presupposition, his strategy is not more invalidate or radical than Damasio's long speculations about feelings as mental images about bodily responses. They both presume more than can be empirically verified. The real question is, at least in the present context concerning the neurophysiological dimension of human affectivity, which research strategy is best at explaining the actual human feelings and emotions[26]. In my opinion, Panksepp lays a better foundation for an interdisciplinary approach to human affectivity because of his insistence on the feeling aspect of emotions. He does not drive a methodological watershed in between emotions

[26] Panksepp is critical of the enthusiasm generated by the newly acquired brain imaging techniques, since the results are often based on a fragile and 'blind' experiential data (what are we looking for and how do we decipher the images of the living brain?): 'Cross-species, experimental analysis, where key brain variables are evaluated, provides an approach that is not full of false negatives and misleading "neuro-echoes" as is modern fMRI [Functioning Magnetic Resonance Imaging, which is based on a combination of registration of neural activity and computer simulation that together generate an image of the brain in function]'. The interweaving of animal neurobehavioral and human neuropsychological research permits substantive dynamic analyses of the core emotions and their associated feelings' (2005a: 26).

and feelings, but emphasize the latter as the most viable access to the first. His dual-aspect monism include evolutionary accounts without excluding feelings as merely subjective 'frills that have added icing to the emotional cake' cooked up by a more important evolutionary process (see note 25). Feelings cannot be methodologically severed from emotions, but their actual presence in human life reminds us that emotional processes are to be considered as unified experiential states, and that these states rely on both objectively detectable evolutionary mechanisms shared by all mammals *and* subjective regulation and refinement due to ethological differences in brain structure (e.g. the neocortex in human) and social interaction. Affective neuroscience does not lock human feelings on to the body; on the contrary, it seeks to clarify the neurophysiological foundation for 'the full emotional feeling state that ramifies throughout the organism' by considering the core affective systems that 'can be regulated, but not created, by higher cortico-cognitive activities' (2005b: 64). Panksepp does not speculate in how humans actually regulate the core affective processes, neither does he pretend that neuroscientific research can effectively clarify the abundant and unique variety of human emotional life (1998: 42). He just claims that human feelings are tethered to the primary affective consciousness of the mammalian brain and argues that a neurophysiological elucidation of the systems operating in this affective consciousness will help understand some of the aspects of the human condition which is, like that of other animals, deeply embedded in affectivity (2004b: 174).

Now, what is exactly Panksepp's proposal for a primal affective consciousness that may help explain the affective nature of human consciousness? He operates with the acronym SELF (Simple Ego-type Life Form) that is embedded deeply within the brain (1998: 309; 2004a: 50; 2005c:178-9). This is a kind of primordial "self-representation" or "self-schema" that originated in the brain stem and developed from the organized motor processes in the midbrain. It is primordial in the sense that it does not at all account for what we know as human consciousness and, furthermore, because it is largely independent of higher cortico-cognitive processes. It operates as an 'archaic SELF-Representation network' (1998: 309) that generates basic instinctual behavior in all mammals. The dynamics internal to this network constitute an affective consciousness that reacts directly to environmental stimuli by initiating endogenous sensory and emotional responses within the subcortical areas of the brain. The reverberations caused by this affective relation to the environment are critical for 'the most crucial biological values that all mammals share' (1998: 183). The mammalian brain is genetically geared to respond in certain ways to the environment, and the intrinsic nature of these response-systems influences the more complex being and behavior of

humans. The instinctual nature of the SELF is structured by various emotional and motivational operating systems that encode intrinsic neurodynamic value structures. These value structures are then experienced by the living subject as raw feelings of pleasure and pain that provide the subject with instinctual choices of responding to the stimuli of the environment. It may seem paradoxical to use the term experience when speaking of instinctual behavior or talking about instinctual choices, but Panksepp's thesis is that the SELF experiences internal raw *feelings* about what is the most appropriate choice or response without having to *think* about it. It is an affective consciousness that feels the value of certain stimuli and thereby enacts a certain reaction. The behavior, however, is not reflective but instinctive in nature, i.e., the system responds immediately without delay due to cognitive elaboration.

In fact, Panksepp operates with different levels of consciousness developed throughout the evolution of the mammalian brain (2004a: 49): first there was *primary process consciousness* that experiences (feels) raw sensory, perceptual, emotional, and motivational properties (instinctual behavior). Then the *secondary consciousness* that possesses the capacity to think about those experiences (non-linguistically). And finally, the *tertiary form of consciousness* that we know as human consciousness and its specific capacity to have thoughts about thoughts. The SELF is generated with the primary process consciousness and reacts to intrinsic values independent of other, more developed, layers of consciousness: 'Our subneocortical animalian brain, with it's many basic attentional, emotion and motivational systems, may actually lie at the center of our mental universe' (idem). What is important in this working hypothesis is not the specific layers (that might seem somewhat artificial), but the general idea that human brains have 'a multidimensional conscious sense of self' (1998: 300) that is rooted in a primordial affective sense of self and the relation to the environment.

Over the last decade, Panksepp has individuated and refined seven distinct endogenous emotional systems in the subneocortical structures of the mammalian SELF (1998: 125-298; 2006: 777-80). These systems all exhibit a certain 'affect logic' (2003: 5) that prevails in human cognitive deliberations and which can be envisioned in basic instinctual animal behavior. As he writes, '[m]y main point is that affective feelings are, to a substantial degree, distinct neurobiological processes in terms of anatomical, neurochemical, and various functional criteria, including peripheral bodily interactions. Emotional and motivational feelings are unique experientially valenced "state spaces" that help organisms make cognitive choices – e.g., to find food when hungry, water when thirsty, warmth when cold, and companionship when lonely or lusty' (2003: 6). The enormous variety of

human affective feelings has, at its core, raw affective experiences that 'appear to be pre-propositional gifts of nature' (2005c:169).

The seven systems that generate the affective internal 'state space' are: SEEKING, FEAR, RAGE, LUST, CARE, PLAY, and PANIC (the systems are capitalized to emphasize the emotional import). Affective experience is satiated with intrinsic values pertaining to these basic systems. The systems function without cognitive interference and can therefore by investigated by a cross-species approach. Here I will not explain the mechanisms involved in all the systems but only briefly outline the main dynamics of SEEKING. The dynamics of the other systems largely reflect those involved in this one.

SEEKING is characterized by a certain feeling tone that 'leads organisms to eagerly pursue the fruits of their environment – from nuts to knowledge, so to speak' (1998:145). It mediates all the appetitive desires that drive and motivate the animal to engage with the environment. It is a self-stimulating system in the sense that the animal is not a passive recipient with regard to the stimuli from the environment, but does actually voluntarily regulate the activity that these stimuli produce in the system. When, for example, the appetite for food is satisfied, the system inhibits the intensity of the stimuli concerning nutrition (1998: 146-7). This is view opposes the classical behaviorist view on animal behavior, since it presupposes an evaluative affective experience that does not fit with a 'cold' mechanistic conception of pleasure and reward. The animal actually feels the raw values pertaining to the objects or events in the environment and thereby establishes an interactive relationship to the objects based on these felt values. Panksepp can, therefore, talk of 'state spaces' in primary process consciousness. The system accounts for the animal's relation to and position in the environment as that of a valenced space experienced by the animal as affective feelings laden with certain values due to the internal structure and dynamics of the system. The same mechanisms are at stake in the other systems, and together the different systems structure the primordial encounter with the world as an affective consciousness that registrates and reacts to certain biological values. In this view, 'affect is an organically embodied part of subcortical instinctual-emotional systems that arouse basic *action-to-perception* processes' (2001: 136). Perception is determined by the affective states of the being that perceives.

In humans, the values experienced as raw feelings have a 'profound influence over what we think and do' (2004b: 177), because the affective values generated in basic system shape and influence the perception of ourselves and the environment (1998: 160). The systems of primal affective and motivational processes that we share with other animals may well be foundational in the evolution

of many of the cognitive skills that characterize human existence (2003: 8). We cannot, according to Panksepp's general idea, approach the feeling aspect of our mental life with a methodological separation of the neurophysiological properties of bodily states and those pertaining to the particular human neocortical capacities. Feelings are not more subjective than emotions simply because emotions are feelings – although these may not be articulated as conscious affordances. The unconscious feelings need not be thematized in order to influence human behavior. We may behave or feel in certain ways that are cognitively impenetrable. On the other hand, not all human feelings are determined by these core emotional systems, although they may find some of their energy or bodily feeling in more raw feeling-states. The individuation of subneocortical systems of core feelings does not exclude the peculiar and profound nature of psychic feelings (1998: 329-30).

Some Considerations on Evolution, Neurophysiology, and Intentionality

Emotions and feelings are a privileged entry into the question of naturalization. Surely, cognition is a biological phenomenon as well as emotions and feelings, but because of the thoroughly embodied nature of emotions they emphasize the inevitable encounter of biological and philosophical investigations of human nature. Damasio and Panksepp present two different perspectives on the neurophysiological dimension of emotions, but they both stress the evolutionary importance of emotional experience. And the theory of evolution indeed seems an unavoidable issue when talking about emotions; in fact, it is not an optional but a necessary factor to take into account (Searle 2000: 40). On the other hand, as Solomon writes concerning an emotion that many evolutionary explanations center on, namely anger: 'There is no doubt that anger (and some other emotions) are part of our evolutionary heritage and include physiological responses that we share with other animals. But this is surely just a piece of the story' (Solomon 2007: 14). Solomon insists that emotions are not only caused by distant evolutionary forces. Our relation to the objects or events of our emotions is also constituted by more than the ancient 'selfish genes' of our ancestors. Emotions are intentional. They contain our most private wishes, desires, and ideas, and thus our emotional relation to objects or events is qualified by this intentionality.

How do we then balance the evolutionary and the intentional aspects of emotions? First of all, we have to distinguish between evolution and biology. Biological explanations are not identical with evolutionary explanations. Biology does not necessarily involve teleological presumptions about the origin of the organisms that it investigates. It is mainly interested in discovering the laws and mechanisms that govern the manifold of biological entities in the world. Naturally, etiological

questions pose themselves more forcibly once we discover 'possible copying mechanisms for the genetic material' (DNA) in all living organisms (Watson and Crick 1952: 737). Evolutionary explanations cannot, however, rely on biology as a justification for its theories. Biology has not yet emptied the mysteries of the living organism; some things still just seem to happen without any explanation (de Sousa 1987: 104-5). This brings us to a second, more interesting, point about evolution and intentionality, namely the question of teleology. How do we account for the objects of emotions? If emotions are not to be considered merely epiphenomena then they must have a reason and enact an actual change in the world. The question is where to look for this reason and change: in the life of the individual person or in the general, anonymous, need for survival. The philosopher Ronald de Sousa has dealt systematically with this problem in terms of the *remote* or *proximate* explanation of emotions (de Sousa 1987: 78). The neurophysiological explanations of emotions are, strictly speaking, proximal; they aim at uncover the neurophysiological dynamics and properties of emotions. Evolutionary explanations, on the other hand, are explicitly remote in that they seek to expound emotions as mechanisms in a general process of adaptation to the need of survival. Now, de Sousa quickly dismisses the simplistic idea of adaptation as survival of the fittest by stressing that 'evolutionary change involves no progress, no inherent direction or "orthogenesis," no built-in drive to mentality or spirituality or group harmony or even complexity' (de Sousa 1987: 80). However, if we subscribe to the general idea of evolution (as de Sousa does and I do too), emotions cannot be completely maladaptive, since the creatures that have them remain alive and even flourish. Emotions do have a biological function that must be envisaged (de Sousa 1987: 195). So to maintain the complexity of human emotions under the pressure of reductive evolutionary explanations, de Sousa looks at the motivational force of emotions; and more specifically, the relation between the biological need and the psychological motivation or want. Emotions influence human behavior at different levels that may be individuated by the notions of instinct and intentionality. First, he distinguishes between what he calls T-instincts and H-instincts (de Sousa 1987: 84)[27]. T-instincts are those we find in simple animal behavior that exhibits predictable stereotyped responses to precise stimuli. H-instincts are human emotions that are experienced as motivational forces, which do not, however, result in fixed patterns of predictable behavior. Now, it is not difficult to point out radical differences of the two types of instincts. It is, however, hard to show how they relate. De Sousa does this by individuating different degrees of intentionality by means of the concept of fungibility. The teleological differences involved in T- and H- instincts are

[27] They are called T-instincts after the Dutch ethologist and ornithologist Nikolaas Tinbergen who individuated this type of basic instincts. H-instincts are particularly human instincts.

determined by the fact that humans have a capacity for singular reference whereas other animals operate with fungible references. To explain the difference, he uses the example of the dog Fido (de Sousa 1987: 98-100) that likes its owner and recognizes this person by means of the physical qualities (smell, voice tone, characteristic movements, etc.). Therefore, we spontaneously believe that Fido actually likes that particular person and not somebody else. This, however, is a mistake. If we imagine that the person died but, miraculously, a perfect clone of that person could be made, then Fido would never be able to tell the difference. He would like that person just as much. The same would not be the case with humans. We know that if we loose a loved one, nobody can ever fill the place of that person. We know that he or she is dead, so if we encounter a person completely identical, we still know that this is not our loved one; he or she is another person albeit the physical identity. This is due to our capacity for singular reference. Humans operate with non-fungible references to their emotional states (I like *this* wallet and *not* another identical one), whereas other animals only refer to the environment in terms of generality (de Sousa 1987: 97). Animals easily transfer the emotion to other objects that have the same general, but not particular, features (e.g., if a dish is empty, the dog immediately turns to another; or, the crows mistake a scare-crows for a real person). The transition from one kind of instinct to another is not due to an ontogentical difference between humans and other animals. It is due to a difference in intentionality. De Sousa talks about a quasi-intentionality in animals that relies on the general character of the reference by which they live their emotional life. Humans are characterized by a 'mental grade of intentionality' (idem) that relies on the specific human logico-linguistic capacities.

What is important in de Sousa's picture of human emotions is that the mental grade of intentionality particular to human beings is not radically irreducible to the common biological story that we tell about other living creatures. Humans are moved by emotions just like other animals and are therefore restricted by some of the same evolutionary forces that are more easily discovered in animals, because their behavior is restricted by the generality of T-instincts. The emotional behavior of other animals are better explained by the remote explanations of evolution, because the general features of their reference (food is nutrition and sex is reproduction). And even though the same forces are present in human behavior, human emotional behavior is different and more complex basically due to the fact that 'we care about individuals' (de Sousa 1987: 100). Evolutionary explanations are important if we want to understand the embodied, physiological nature of human emotions. We are rooted in biology particularly by our emotions and to fully understand these ubiquitous phenomena in human life, we need to clarify the neurophysiology operating in them. The

reason why an evolutionary account is never enough is due to the singular reference of human emotions. Human emotional behavior relies on a proximal explanation. As I mentioned above, the neurophysiological explanation is carried out at a proximal level. However, the explanation needs to be proximal in another sense. Contrary to other organisms, humans have a *biography* and their desires are *time-indexed*, therefore 'having a causal history is not sufficient' (idem). Human emotions cannot be understood if we do not take into account the specific history of the individual person (de Sousa 1987: 104-5). Why I scream and cry over having lost an old worthless wallet is perhaps because that wallet has accompanied me a long time, or because it was a gift from a person that I care about[28]. The idea of personal history and time will have important bearings on the following; but before turning to my own account of human affectivity, I will briefly try to position Ricoeur in relation what has been said about the neurophysiological dimension of emotions.

Conatus: Spinoza, Ricoeur, and Neurophysiology

Although Damasio and Panksepp both adhere to evolutionary explanations of human emotions, it should be clear from what has been said above that their neurophysiological explorations do not coincide. Damasio consider human feelings to be perceptions of bodily changes that are the real emotions, and as such feelings are mere subjective registrations of more fundamental and objective emotions. We have to understand the specific human feelings by clarifying the neurphysiological dynamics of bodily emotions (Damasio 2004: 56). Panksepp, on the contrary, resists this methodological separation of feeling and emotions, which he sustains leads to an ontological neglect of basic affectivity (Panksepp 2005c: 160). Emotions are feelings, and that already on a pure bodily level. Consciousness is affective consciousness, in humans as well as in other animals. The cognitive capacities in humans is not 'added-on' to the body, but deeply immersed in the affectivity that characterizes the whole organism. We do not perceive or read-out the emotions and therefore feel in a certain way. On the contrary, affective feelings shape and influence our perception and cognition in such a way that our higher brain functions are always in a certain affective state. Therefore Panksepp writes: 'Descartes's faith in his assertion "I think, therefore I am" may be superseded by a more primitive affirmation that is part of the genetic makeup of all mammals: "I feel, therefore I am"' (Panksepp 1998: 309).

[28] The temporal and personal aspect of inanimate things is a hotly debated topic in contemporary social psychology, which goes under the name 'endowment effect'. People tend to value things higher once they are actually endowed with them (see Goldie 2007: 111).

Despite their differences, Damasio and Panksepp nevertheless resort to the same thinker as the speculative background for their empirical research, namely the Dutch seventeen century philosopher Baruch Spinoza. This is not uncommon among philosophically interested neuroscientists scientists; for example, the neurobiologist Jean-Pierre Changeux confesses, in a extended discussion with Ricoeur, his affection for Spinoza's monism and that he himself is striving for a new synthesis of the objective science and subjective experience (WT: 8/14, 27/35-6, 31/39). What is interesting here is that Ricoeur himself uses the Spinozian concept of *conatus* as a key-concept in his theory of subjectivity, and particularly with regard to human affectivity. Ricoeur transcribes conatus as the originating affirmation or attestation (in the later works), that is, the subject's basic 'will-to-live' (vouloir-vivre) or original will-to-say-yes. The concept of an originating affirmation is more than just a will to survive and yet even more basic than rationality (TFA 290-1/361-362). It is a primordial energy that drives the subject to attest its being in the world (FP 46/53). At a first glance, it might seem a wild metaphysical postulate without any serious grounding or possible argument. But if we look at the way that neuroscientists employ this Spinozian heritage, we might nourish it a bit with some plausible arguments.

Damasio is the neuroscientist who has treated the heritage from Spinoza most carefully. In fact, his last book is partly dedicated to demonstrating how Spinoza's philosophy would help us understand the biology of the mind (2003: 13). He is particularly interested in monistic idea that the mind and the body are two attributes of the same substance, and furthermore, that the human mind is the idea of the human body (2003: 12). But most of all, it is the notion of conatus that attracts his attention because, according to him, it can be translated directly into current biological terms of genetically determined dispositions in the brain that seeks both survival and well-being (2003:36-7). The conatus goes well with his somatic-marker hypothesis in the sense that the conatus renders well the emotional forces present at 'the theatre of the body', i.e., the manifold of bodily responses, that highlight the positive and negative values in the experience of the world. All living organisms endeavor to preserve themselves and, furthermore, optimize pleasure and minimize pain. The conative forces are not conscious but work with the anonymous energy of natural wisdom that is the product of millions and millions of years of evolution. The conscious, subjective experience of these anonymous forces is described in the following way: 'When the consequences of such natural wisdom are mapped back in the central nervous system, subcortically and cortically, the result is feelings, the foundational component of our minds' (2003: 79). This conception of conatus as an aggregate of evolutionary forces that automatically drive the organism toward bodily self-

perseverance and satisfaction, which is later extended with a cognitive brain map and mental ideas of these reactions, the so-called 'machinery of feeling' (2003: 80), presents a somewhat distorting and poor interpretation of the Spinoza's use of the concept. The hypothesis that the mind is the idea of the body does not entail that human knowledge is completely tethered to that body. On the contrary, Spinoza writes clearly in his *Ethics*: 'Body cannot determine mind to think, neither can mind determine body to motion or rest or any state different from these, if such there be' (Spinoza 1955: 31; III, pro. 2). Now, I do not want to go into the interpretive difficulties of Spinoza, but only emphasize that the Spinozian monism does not necessarily lead to the conclusion that mental ideas always recur to bodily states. The conatus is a general endeavor or will-to-live that is irreducible to the more restricted idea of evolutionary forces. This general will-to-live is surely influenced by the process of evolution, but there is more to the notion, when we consider human affectivity. Ricoeur points to this in his discussion with Changeux. One has to read Spinoza's *Ethics* more carefully when talking about human affectivity and the nature of the conatus. There is a different finality at stake in human endeavor to preserve its life simply because there is more to human life than the absence of death. Ricoeur agrees to the idea that conatus is indeed '[t]he effort to live – the desire to exist' (WT 229/242) but retain that we have to see this effort in relation to the idea of 'the good life' (WT 228/241). This is somewhat similar to de Sousa's analyses of remote and proximate explanations. In order to fully understand how the conatus works in the human subject, we have to take into account both the evolutionary and personal aspect of the subjective will-to-live. We may have the same anonymous forces working within us, but these are always experienced in relation to our personal characteristics or biography (Ricoeur would say our 'existential difference'). The conatus that drives the human subject relies on, at least, both these factors.

Now, in my opinion, Panksepp's use of Spinozian monism is more congenial to Ricoeur's idea of the conatus. Like Damasio, Panksepp subscribes to the idea that mind and body are two attributes of the same substance, namely his dual-aspect monism (Panksepp 2007: 232; 2005a: 23; 2005c: 163), and he points to this idea as Spinozian (2004b: 41-2). The neurobiological and psychological properties of the brain are of the same substance, but they have to be understood from different perspectives. The neurophysiological investigation must focus on the basic affectivity that is common to all mammals and which results in '*basic* emotional tendencies' (2006: 775). Such an investigation might clarify the 'ethological animal brain' and thus 'we should be able to shed light into core emotional tendencies by studying animal brains' (idem). He focuses on the 'intrinsic value guides for existence' (2005a: 22) that characterize all animal affectivity and the consequent

behavior. However, his methodological starting point (all mammals have affective feelings) allows for what he calls 'conative variants' (2005b: 40) in mammalian affectivity due to the variation of brain structures in the different animals. The conatus is not the same in humans as in other animals. We may share many of the basic instinctual feelings, but there is something more to consider in human feelings. Such a neurophysiological approach seems to go better with the philosophical treatment of the conatus in Ricoeur.

Ricoeur defines affectivity in the following manner: 'In a general way, affectivity is the non-transparent aspect of the Cogito. We are right in saying "of the Cogito." Affectivity is still a mode of thought in the widest sense. To feel is still to think, though feeling no longer represents objectivity, but rather reveals existence […] We can express it otherwise by saying that through feeling the personal body belongs to the subjectivity of the Cogito' (FN 86/83; cf. FM 131/147). Feelings are thoughts; not a disinterested categorization of objects, but the revelation of existence. Our thoughts about the world, the other, and ourselves are qualified by feelings in the sense that affectivity 'personalizes reason' (FM 102/118). The environment becomes a world for me; a world characterized by heterogeneous values that orient my behavior. The conatus, my will-to-live, drives me to engage myself in my existence in this world, but my affirmation is always affected by the values that I experience in my affective relation to the world and my own engagement. Now, if we consider Panksepp's neurophysiological model of affectivity as the experience of raw emotional values that characterize basic mammalian instinctual behavior, we might make Ricoeur's notion of conatus as originating affirmation more than just a speculative postulate. Panksepp's empirical studies argue for a basic mammalian affectivity that influences human behavior, because our human feelings are deeply rooted in a common ethological ground. Our behavior is not qualitatively different from the instinctive conduct of other animals, but, as Ricoeur writes, '[i]t is the instinctive conduct which decreases in the case of man. Man has quantitatively more instincts, if we include the new anxieties and new incentives which he invents, but he is less instinctive if we stress his loss of unlearned forms of conduct, spontaneously adapted from his milieu' (FN 95/91).

Ricoeur emphasizes the importance of basic instincts in human behavior, but insists on the increased quantity of instincts involved in human life. The human desire or will-to-live is different from that of other animals because of the mediated nature of the relation to the environment and to itself as a human subject. The higher cognitive human capacities that enable this mediate relation are indeed bound to the ecological life of the subject, which 'attest to the nonautonomy of knowledge, its rootedness in existence […] Thereby is discovered not only the unsurpassable nature

of life, but the interference of desire with intentionality, upon which desire inflicts an invincible obscurity, an ineluctable partiality' (FP 458/442). In human emotional experience there always remains 'an invincible obscurity' due to, at least partly, the ecological nature of human life. We cannot understand human emotions without resorting to general ethological accounts that often include evolutionary explanations. Many of our emotions are driven by desires whose immediate satisfaction is rendered impossible by the multitude of laws and rules of conduct instantiated throughout the process of civilization (Goldie 2000b: 33). I might, for example, intensely want to strike a blow at the idiotic college who has insulted me; instead I yell at him and thump my cup down on the table and leave the room. I have momentarily vented my anger, but I may remain with a feeling that a fist in his face would have been a more appropriate expression of my anger. It would not, however, be the most appropriate in the given context, since I would properly loose my job and perhaps even face a law suit. These immediate reactions might find their explanation in some basic affectivity that generates instinctual mammalian behavior. Human emotions, however, contain more than these basic affective feelings. And we cannot explain them only with an archeological survey into the mammalian brain systems. Ricoeur therefore calls out for a dialectics of archeology and teleology as the adequate approach to human emotional experience (above, pp. 62-5). The complex nature of human emotions is related to the larger quantity of human instincts generated by the symbolic nature of our interaction with the world and other human subjects. Our emotional behavior is influenced by the peculiar, proximate teleology that governs human existence, namely the inscription in a society with other human subjects. And further, we do not just want to live among other subjects; we want to be recognized as a self by other selves (in our pursuit of 'the good life', of happiness). This teleology of recognition is indeed rooted in and, to a certain degree, shaped by the organic nature of our being, but somehow it rebels against the organic nature of life. Ricoeur writes: 'No doubt the passion to achieve recognition goes beyond the animal struggle for self-preservation or domination; the concept of recognition is not a struggle for life; it is a struggle to tear from the other an avowal, an attestation, a proof that I am an autonomous self-consciousness, But this struggle for recognition is a struggle in life against life – by life […] This is the sense in which desire is both surpassed and unsurpassable. The positing of desire is mediated, not eradicated; it is not a sphere that we could lay aside, annul, annihilate' (FP 471-2/455-6).

Ricoeur's use of the Spinozian concept of conatus is not contrary to or hostile toward a neurophysiological approach grounded in an evolutionary theory. On the contrary, Ricoeur emphasizes the basic mammalian affective feelings present in human emotional experience (e.g. the

struggle for survival and the seeking of pleasure and avoidance of pain). What he does contest, however, is the reduction of human emotions to the regulation of physiological responses in the body. There is another kind of teleology at play in human emotional experience, namely the desire for recognition. We want to be part of a human society as well as staying alive. Our emotional behavior is therefore conditioned and articulated by our existence in that society alongside with the organic values of our biological life. Furthermore, humans have the capacity for singular reference which means that the objects of human emotions often are of a non-fungible nature. We care about singular object and individuals, and because of this our feelings are much more differentiated and complex than those of other animals. Some of the obscure and seemly irrational features of emotional experience may very well find their explanation in the evolutionary forces present in our basic affective relation to the environment. And a clarification of the neurophysiology of emotions may help us understand why we sometimes feel and act irrationally. The body has its own reasons that deeply influence our being in the world, and therefore one should not underestimate the embodied nature of emotions. But, as Solomon rightly pointed out, this is only part of the story. Although we might be driven by the same conatus as other animals, there are important differences ('conative variants' as Panksepp call them) when it comes to human affectivity. These variants or differences are mainly due to the higher cognitive skills present in humans, and which might be characterized by what de Sousa names 'full-fledged mental intentionality' together with the fact that humans have a biography and time-indexed desires (de Sousa 1987: 100).

In the concluding sections, I will try to present an approach to human affectivity that account for the difficulties discussed above. It will only be an outline of theory that I have argued for in more detail elsewhere[29]. What is important in the present context is to show why an articulation of human affectivity is important for the relation between subjectivity and ethics.

[29] I am presently (from 2005) working together with the psychiatrist Giovanni Stanghellini on a theory of emotions that seeks to combine a philosophical and psychopathological approach to human emotional experience. We have written a forthcoming article (Rosfort and Stanghellini, forthcoming) and are currently working on other three articles about the Heideggerian concept of Befindlichkeit, the historical implications of the concept of mood, and the relation between personhood and the embodied nature of emotions. These articles are going to be the groundwork for a co-authored book (Emotions and Personhood: A Philosophical and Psychopathological Exploration). I owe many of the following arguments from this collaboration, both from long discussions at his house, in trains, at his seminars, at the Department of Psychology at The University of Chieti, Italy, and finally from his books (Stanghellini 1992, 1997, 2004, 2007). I have chosen not to include his extensive work on psychopathology in the following analyses, because this would make the arguments more difficult and therefore more the presentation less poignant. However, my treatment of moods and affect draws substantially on arguments from our forthcoming article mentioned above.

Emotions and Personhood: Moods and Affects.

It is time to narrow down the concept of affectivity and look more closely on what is peculiar about human affectivity. De Sousa has recently observed that '[b]eyond the 'core systems' described by Panksepp, the set of emotions that each of us is capable of experiencing are, like the belief we hold, both potentially infinite and, in their totality, unique to each individual' (de Sousa 2007b: 329). As we have seen, human affectivity has much in common with that of other animals; nonetheless, the behavioral neuroscientist T. M. Preuss calls for caution when we observe ethological similarities among mammals because '[e]ven ignoring our intellectual abilities and cultural accretions, people are most peculiar beasts' (Preuss 2004: 5). And the main reason why human affectivity is drastically different is, because human beings are persons other than being part of a phylogenical staircase. The concept of emotion and the concept of a person have to be considered as parts of the same experience. Human emotional experience cannot be understood without considering the nature of the being that has this particular experience. This means that the conception one has of the person will influence one's view of emotions, and vice versa. Although emotions occur in most animals, some emotions are particularly human such as shame, resentment, delusion, pride, sadness, love, guilt, hopelessness, euphoria, and so on. As Charles Taylor observes: 'The peculiarity of these emotions is that it is at one and the same time the case that our formulations are constitutive of the emotion, *and* that these formulations can be right or wrong' (1981: 101). Especially when interpersonal emotions (such as shame or guilt) are concerned, what they are to a person depends significantly on what that person takes them to be. Not only is our understanding of the emotions that we experience constitutive of the emotion itself, but also our understanding of ourselves is constitutive of the emotion. Furthermore, if our understanding of an emotion can be right or wrong, then this entails that emotions tell us something about ourselves as persons, about our situation and our behavior (and its motivations). But now I will first briefly define some of the fundamental aspects of personhood as I see it.

Personhood

I concluded above that Damasio's conception of the self is oversimplified (although he claims it not to be the case (Rudrauf and Damasio 2005: 239)), because it reduces the self to a living system in time and space primarily controlled by what Harry Frankfurt once called 'first-order desires' or 'wantons' (1971). Our mind is thus reduced to a complex survival machine, intrinsically elaborated during evolution, wherein our body highlights the most favorable options in the process of our

decision-making. However, there is more to the sense of self than just perseverance or immediate desires. Indeed, we have to *want* to be the self that our organism or other people tell us that we are. Frankfurt points, once again, to the fact that humans are peculiar beasts: "They are capable of wanting to be different, in their preferences and purposes, from what they are […] No animal other than man, however, is capable to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires" (1971: 7). To be a person involves the capability for evaluating the identity and change that constitute one's temporal and situated self.

In virtue of our reflective self-evaluation, we distance ourselves from our immediate desires and inclinations, and might perhaps choose something even contrary to those because of another desire or a sense of duty, shaped, for instance, by imagination or emanated from a firm principle. Thereby, we create a second-order desire that is completely personal and may even appear idiotic in the eyes of other people. For example, a person who commits suicide revolts against an organic inclination to live, against all possible rational arguments about the incomparable value of life and perhaps even against her or his own desire to live. We can be angry, condemn the act, remain without an explanation ('she seemed to have all the heart can desire'), but it would never cross our minds to contest her personhood because of her decision. In fact, Ricoeur argued that the autonomy of decision is a structural capacity of the subject, because we are able to 'designate in emptiness' (désigner à vide); the human subject is structurally pure possibility, although some possibilities might be absurd (above pp. 29, FN 43/42). Thus, to be a person is not essentially to be rational, but *to have a will* (Frankfurt 1971: 11; cf. Blackburn 1998: 65). We are persons because we are capable to choose to be what we want to be. Although this may seem an obscure reminiscence of existential philosophy that goes against intuition and the laws of bio-physics (for example, a man who want to be a woman because he *feels* that he is a woman captured in a man's body), this capacity still prevails as a fundamental characteristic of being a person[30].

A person is a contextualized and situated self with intentional attitudes, characterized by ontological fragility (above, pp. 50-2) and capable of self-evaluation, as opposed to a minimal self (Gallagher and Marcel 1999; Zahavi 2005) or a core self (Damasio 1994, 1999; Russell 2005). The reason for this is that '[der] Mensch ist als Person immer schon ein 'homo duplex'. Er steht sich selbst gegenüber und wird sich fragwürdig; er steht auf eine Bühne vor seinesgleichen und schwankt zwischen Sein und Schein' (Fuchs 2002b: 144). The term 'homo duplex' was originally

[30] As Goldie points out (referring to Augustin): 'He wants to *want* to be other than the way he is' (2004: 109).

elaborated by the 19th century French philosopher Maine de Brian[31] in a context wherein man is seen as a single nature in his vital states of being but as a double (or complex) nature in belonging to humanity. We are passive in our nature as sentient beings because we are shaped by our physiological nature, environment, and past history. Yet, we are active in virtue of being persons. On the one hand, we are a part of nature, and therefore nature sets a pre-reflective agenda for our behavior in the form of instincts and emotions (an affective core SELF endowed with basic instinctual behavior); in short, an extremely complex machine set to solve the problems encountered during our lifetime in the ongoing interaction with the environment. On the other hand, as Ricoeur points out, '[m]ore fundamentally man is a being who poses problems and raises questions – if only by bringing into question the very foundations of that society that bids him to adapt himself quietly to its systems of work, property, law, leisure and culture' (FM 101/117; cf. FN 227-9/213-4). Our *being* this 'double' person, endowed with a specific physical constitution and situated in a certain historical, social, and cultural context, is complicated even further when we consider our capacity to transcend our being this person by questioning it. Not only do we question our being a part of nature and the social and cultural context wherein we are situated, we also question our own identity as being this particular person, constituted by a certain physical constitution, past history, present environment, and character.

Being this specific person might be a problem to me (remember that Ricoeur characterizes personhood as a task, not a fact (FM 61/86)). I may *feel* that I am different from what I appear to be. This often begins with a diffuse but pertinent feeling, which slowly develops into a conviction that the person I appear to be does not express the person who I really am. We might say that the person is situated in the dynamic dialectic of 'being' (*Sein*) and 'appearing-to-be' (*Schein*)[32]. Because of the fundamental possibility of wanting to be different, we always evaluate our being and our

[31] The full sentence goes 'Homo simplex in vitalitate, duplex in humanitate' and is, as mentioned, quoted several times by Ricoeur (e.g. FN 228/213, 249/234; FM 91/107) to recall the dramatic dualism that we experience as a daily reality.

[32] Nietzsche and Kierkegaard have dedicated much of their work to this problem. Kierkegaard writes eloquently: 'Personhood is a synthesis of possibility and necessity. Its continued existence is like breathing (*re*spiration), which is an inhaling and exhaling. The self of the determinist cannot breathe, for it is impossible to breathe necessity exclusively, because that would utterly suffocate a person's self' (Kierkegaard 1983: 40/172). The philosopher Arne Grøn has worked much on this theme in Kierkegaard (e.g. 2004a): 'In seeking to account for the embodied and embedded nature of human consciousness we have in various ways encountered a remarkable feature of the human being, a double character or a *redoubling*: to exist is to be in a process of becoming *and* to relate to oneself in this process' (34); or said otherwise: 'we are heterogeneous to ourselves, and our identity as unity is a problem to ourselves' (43n). Nietzsche, on the other hand, has emphasized the interesting dialectic of 'being' and 'appearing-to-be' in the development of the self, e.g.: '*How appearance becomes being* […] If someone obstinately and for a long time wants to *appear* something it is in the end hard for him to *be* anything else' (Nietzsche 1986/1954a: 39/487, §51); or as he writes elsewhere: 'What are our experiences, then? Much more what we attribute to them than what they really are. Or should we perhaps say that nothing is contained in them? that experiences in themselves are merely works of fancy?' (Nietzsche 2007/1954b: 128/1096, §119).

appearing-to-be, and, in some way or another, relate ourselves to ourselves in this dialectic. To be a person is a normative matter.

Feelings of Being-in-the-World

The fundamental way of being situated in the world is, before it becomes an explicit theme for rational thinking, being in a given feeling or emotional tonality. One approach to the normative character of personhood is to investigate into this implicit feeling or emotional tonality of human existence; this entails, as we will see, looking at the workings of moods and affects.

Feelings are the primordial medium in which I encounter the world as a set of *affordances*: a set of relevant possibilities that are my own possibilities as an embodied person situated in this particular world. This being situated in a world of possibilities through certain feelings is described by Heidegger as *Befindlichkeit* (1996: 126-131/134-40); a term that combines the notion of location (finding oneself somewhere) and that of being in a certain feeling or attunement (the English translation of the German *Stimmung*). It is equally wrong to say that a certain thing in the world causes a specific feeling (e.g. that a wild animal causes fear) or that a certain feeling colors my perception of the world (e.g. anxiety causes me to see a stick as a snake). There is no causal relation between perception and feeling-state, or vice versa. Both feeling and perception find their explicative correlate in the actual and concrete situation of the person. We feel and perceive the world in a certain way because of our being-in-the-world (*In-der-Welt-sein*). The main force of the concept of *Befindlichkeit* might be that it emphasizes the work of feeling as a disclosure (*Erschliessen*) of one's embedded and situated relation to the world and oneself, and not an impediment to objectively appreciate a certain state of things. The significance of an event or state of affairs is not merely a matter of its intrinsic properties, rather of its relation to my concern for and current engagement in the world, the other and myself. Feeling states reveal how the world is *for me*. The concept of *Befindlichkeit* is closely tied to that of understanding (Verstehen), which Heidegger has spent much of his work to argue for. He writes, for example: 'Attunement [Befindlichkeit] always has its understanding, even if only by surpressing it. Understanding is always attuned [gestimmtes]' (Heidegger 1996: 134/142). We can only understand ourselves and the world in which we are situated through the context of our practical engagement, and, as we have seen, this engagement is primordially enveloped in a certain feeling. This way of looking at feeling has fundamental implications for the understanding of emotional experience that will be best

appreciated if we first discuss the relation between feelings and emotions and then that between moods and affects.

Feelings and Emotions

Instead of fighting one another, the feeling theories and the cognitive theories should learn from one another[33]. Emotions are rooted in both physiological reactions and psychological phenomena. On the one hand, 'emotions are bioregulatory reactions that aim at promoting, directly or indirectly, the sort of physiological states that secure not just survival but survival regulated into the range that we, conscious and thinking creatures, identify with well-being' (Damasio 2004: 50). On the other hand, '[o]ne can be angry without feeling anything in particular, without doing anything in particular and without displaying any physiological symptoms of a unique syndrome for that emotion' and thereby the feeling theories 'ignore what might be called *subjectivity*, one's viewpoint and what one experiences – other than sensations and their like – when he or she has an emotion' (Solomon 1977: 44). Although I doubt if feelings can ever be completely absent, Solomon is right, however, in pointing to the fact that an emotion is not always bound to the ebbs and flows of the bioregulatory reactions of our body, but is constituted by subjectivity as well as physiology. The following approach to human affectivity employs important insights from both the cognitive and feeling theories, but differs from these by the emphasis on the concept of personhood as fundamental to human emotional experience.

Emotions are kinetic, dynamic forces that drive us in our ongoing interactions with the environment. This definition of emotion focuses on the embodied nature of emotions, but rejects its reduction of the body to the object-body or physiological mechanism (e.g. visceral changes registered by higher cognitive brain systems). It obviously also rejects the conceptualization of emotions as pure mental phenomena since an emotion is not a purely and primarily cognitive phenomenon affecting the mind, but a phenomenon rooted in one's lived body, and it can to a certain extent be subconscious (Prinz 2005: 15-18; Pankseep 2005b: 61-3). One might say that emotions are characterized by their connection to *motivation* and *movement*. Emotions are functional states that motivate and thereby produce (or hinder certain) movements. As functional

[33] This is, in fact, becoming more frequent in the philosophical research on emotions. As we have seen above (note 19), Solomon has, in his later research, come to consider feelings an important part of the otherwise cognitive emotions. And Goldie, who is emphasizing the feeling aspect of emotions, considers intentionality fundamental in at least some feelings that he calls 'feeling towards' (2000a: 58-62). However, the dichotomy still exists and more work has to be done in order to sort out the complex relation between cognitive and feeling aspects of emotional experience.

states that motivate certain movements (and check others), emotions are closely connected to dispositions in the sense that they dispose the person in a felt readiness or impediment for action.

Feelings make up a crucial part of emotional experience but, as we have seen, there is disagreement about how they do so. James identifies feeling with emotion (1884: 190), and yet Solomon distinguishes the two by saying that we can experience an emotion without feeling. Still, to detach feeling from emotion includes the risk to over-intellectualize emotion (Goldie 2000a: 41). If we consider an emotion structured as a feelingless judgment about the world, then we loose an important criteria of distinction between actions done out of emotion and actions not done out of emotion. For example, a judge should be able to treat a person regardless of how she feels about him. Only thereby can we say that it was a sentence built on just causes and not on the personal feelings of the judge. The feeling dimension of emotion is what permits to distinguish emotion among other cognitive functions (perception, deliberation, evaluation, judging, etc.). However, the distinction between emotion and other cognitive functions is often blurred by the fact that it is difficult to ascertain whether an action is done with or without an emotion. Emotions immerse the other cognitive functions in their complex movement whenever they appear. The judge is perhaps repudiated by the appearance of the accused or by his word. Or she may be in an irritable mood without being able to put her finger on what exactly it is due to. And both of these emotions might influence her final sentence. Of course, a good judge is one who is able to set aside these personal aspects. Nonetheless, the fact that is difficult to be a good judge reveals something important about the feelings of emotions, namely that they tend to impose themselves on all our thoughts and actions. Feelings are a constitutive part of all emotional experience, although how an emotion feels may remain vague to the subject. In order to differentiate among different emotional experiences, we need to take a closer look on the feelings that they cause. Feeling is, in my opinion, the best way to reveal that we are in a certain emotional state.

One reason why the feeling dimension of emotion is important is because emotions can be both conscious and subconscious. Conscious emotions take up, by virtue of their intentional attitude, a substantial part of our attention in a given situation. For example, I choose not to undress myself in front of other people because I feel embarrassed. Unconscious emotions, on the contrary, are not direct objects in our attentional field. They manifest themselves through certain feelings. These feelings can be vague and opaque (as in the case of moods; see below). However, the feelings that an emotion causes are an essential component of the emotion itself, since we, as persons, need to

acknowledge these feelings in order to fully access the emotion[34]. For instance, our good judge may be in a 'bad mood' the day of the trial. This is indeed a very vague constellation of feelings. Her body feels heavier than usual and the sunlight is annoying. The coffee tastes strange, and even the smallest obstacle, such as a binding door, leaves her exasperated. But although the expression 'bad mood' might seem innocent and insignificant it is, however, wrong to say (as Ratcliffe does, 2005: 55) that this is a rather superficial subjective state. Indeed, being in a 'bad mood' often involves profound feelings that reveal oneself as being in a certain dispositional state and readiness to a (although still implicit) given action. We should not let our rather casual use of words obscure the importance of moods, because, as Ryle writes, '[m]ood words are short-term tendency words, but they differ from motive words [i.e. inclinations], not only in the short term of their application, but in their use in characterizing the total 'set' of a person during that short term. Somewhat as the entire ship is cruising south-east, rolling, or vibrating, so the entire person is nervous, serene, or gloomy' (Ryle 1949: 100).

Furthermore, the constellation of feelings involved in an emotional experience may contain more or less explicit cognitive elements. The feelings involved in fear, for example, may block our higher cognitive skills in order to promote the immediate instinct to flight from the object that causes the emotion (e.g. Panksepp's basic affective instincts). In sadness, however, feelings and cognition are more intrinsically intertwined. The feelings are both subject and object for our reflections. We are sad because we feel sad, but the thoughts involved in our sadness may enhance or diminish our feeling of sadness.

And finally, time plays a fundamental role in the feelings involved in an emotional experience. Whereas some feelings are more or less instantaneous (panic, joy, sexual arousal), others are prolonged (grief, hatred, boredom).

In order to access these emotional states we need to pay attention to the diffuse and vague constellation of feelings involved in our interaction with the world. This is not an easy job. Nevertheless, I believe that by sorting out the main characteristics of the different feelings involved in two main emotional states, moods and affects, we gain an effective tool to close up on an understanding of how the person lives, experiences and understands their emotional states. Both moods and affect are characterized by the constellation of feelings involved in the experience of

[34] Prinz argues convincingly for unconscious emotions (2004: 201-5). However, he insists on relating these exclusively to 'core relational themes' regarding basic 'dangers and goal attainment' that once again point to a distant evolutionary teleology or explicit cultural influences. I, on the contrary, shall argue that unconscious emotions such as moods play a much more proximate and subtle role in the person's life.

them, and an analysis of the phenomenology of these feelings enables a better understanding of the emotion, its influence on the person, and, finally, of the nature of personhood itself.

Moods and Affects

Phenomenologists have contributed to explaining this distinction in an explicit and systematic way (for an overview see Strasser 1977, Smith 1986, and Fuchs 2000). This distinction is merely incipient in Husserl's writings and is afterwards made explicit by Scheler (1912; 1916), Heidegger (1996), Sartre (1939), Ricoeur (FN; FM), Stephan Strasser (1977), and Thomas Fuchs (2000): whereas affects are responses to a phenomenon that is grasped as their motivation, moods do not possess such directedness to a motivating object. Although their terminology differs, and that often confusingly (Scheler: Affekten/Gefühlen; Heidegger: Affekten/Stimmungen; Sartre: affects/emotions; Ricoeur: sentiments schematizes/ sentiments informes; Strasser: Emotionen/Stimmungen), their analyses of the phenomena concur in the general characteristics (see Table 1 below).

Affects are focused, intentional, and possess directedness. Affects are felt as motivated; they are more determinate than moods and more articulated. Affects do not open up a horizontal awareness, but occupy all my attentional space (e.g. in fear I am completely absorbed by the phenomenon that terrifies me). When I am affected, a relevant feature of the world captivates me, irrupts into my field of awareness without me having decided to turn my attention to it. I become spellbound to it and all my attention is captured by it. Typical examples of captivating affects are grief (when the death of a beloved person occupies all my attentional space) or phobias.

Moods, on the contrary, are unfocused and non-intentional. They do not possess directedness and aboutness. They are felt as unmotivated, and there are no 'felt causes' for them. They are more indefinite and indeterminate than affects and are often inarticulate. Moods have a horizontal absorption in the sense that they attend to the world as a whole, not focusing on any particular object or situation (cf. Ricoeur's 'fundamental feeling' (FM 105/121)). Moods often manifest themselves as prolonged feeling-states as opposed to the more instantaneous nature of affect[35]. Whereas most affects fill up the whole field of awareness for a brief period (for example in fear or anger), moods convey a constellation of vague feelings that permeate my whole field of awareness, and they often last for a longer period than affects.

[35] In fact, Russell's neuropsychological account of core affect ("the simplest raw (nonreflective) feelings evident in moods and emotions") defines moods as "prolonged core affect" (2003: 147-148).

Moods are global feeling-states that do not focus on any specific object in my field of awareness. When we are in a certain mood we relate ourselves to the world and to our own person *through* that mood (FM 131-2/148). In euphoria the perception of my body is feeble and may even vanish. I feel absorbed in my concerns; my self-awareness, my body, and the world fuse together in perfect harmony. In sadness the perception of my body comes to the foreground; I may feel my body as an obstacle, a hindrance separating me from the world and perhaps even from myself (Fuchs 2002a: 123). Thus, moods are atmospheric and often corporeal in that they permeate my perception of the environment. They can bring me closer to or distance me from the world in that they elicit a certain atmosphere that becomes the tonality through which I perceive the world and myself. When I am feeling happy, the world and other persons appear in a soft light of possibility and openness; I feel differently when I am jealous. In this case, things appear as prowling perils; even the most sincere smile might be perceived as false and dangerous to my person (Smith 1986: 131-4).

An important aspect of moods is that in virtue of being prolonged and pervasive feeling-states, they are dispositional in nature and may develop into character traits: "our traits are shaped by our emotions and moods, just as our emotions and moods are shaped by our traits" (Goldie 2000a: 141). Goldie here puts forward an interesting dialectic that I will return to in the end of this section.

The table below roughly resumes the main characteristics of moods and affects in such a way that their oppositional nature becomes clear.

Table 1[36]

Mood and affect

Mood	**Affect**
Unfocused	Focused
Non-intentional	Intentional
Not motivated	Motivated
Inarticulate	Articulate
Horizontal absorption	No horizontal absorption
Emanated from, not by	Emanated from and by
No captivation	Captivation
No 'felt causes'	'Felt causes'

[36] Adapted, with modifications, from the descriptions in Smith (1986: 109-46). Recently an interesting empirical investigation on how 106 non-academic people distinguished between emotion and mood (here I prefer the term 'affect' to 'emotion', but the characteristics coincide) produced a result very similar to the table presented here. (Beedie, Terry, and Lane 2005: 817).

Indefinite and indeterminate	Determinate
No directedness	Directedness
Sustained	Instantaneous

Examples of moods as opposed to affects are anxiety as opposed to fear, sadness (as opposed to) grief, euphoria/joy, dysphoria/anger, tedium/boredom.

There is little work done on moods in the Anglo-Saxon philosophical tradition. The treatment of moods is often reduced to peripheral mentioning in larger analyses of emotions[37]. One might say that moods are not considered to be a serious philosophical matter, but fairly superficial feeling-states[38]. If they are mentioned, their main characteristics concur with the phenomenological account, namely an objectless phenomenon (for opposing views: Kenny 1963: 60-62; Crane 1998: 238-46; Goldie 2000a: 143-51). However, the fact remains that moods are often treated as opposed to emotion (e.g. Blackburn 1998: 130, de Sousa 1987: 311) so that the distinction is not mood vs. affect, but mood vs. emotion. I believe, however, that if we treat both mood and affects as different constellations of feelings within the emotional experience, then it enhances the possibility to pursue an understanding of the dialectical transitions between the two phenomena and their significance for the person. Whereas affects seem to enjoy a firm definition because of their object-directed nature, moods (due to their disoriented and hazy phenomenality) do not enjoy the same privilege. It can therefore be of help to elaborate more thoroughly on the principal characteristics in the relation to the concept of a person.

A very general definition of my use of the term emotion might be like this. I use emotions as an 'umbrella' term that describes what is known as emotional experience. This experience is characterized by different constellations of feelings that range from brief, clear affects at one end to longer, more diffuse moods at the other. What we experience as different emotions lies within these two extremes and can be categorized according to the nature of the feelings involved (intentionality, temporality, and cognitive elements evidenced by narrative structures). But more about this in the following.

[37] An interesting exception is Gilbert Ryle's succinct analysis of moods in emotional experience in his *The Concept of Mind* (1949: 98-104).

[38] The philosophical treatment of moods often coincides with the attitude of the influential psychologist Paul Ekman: 'I argued that emotions are necessary for our lives, and we wouldn't want to be rid of them. I am far less convinced that moods are of any use to us. Moods may be an unintended consequence of our emotion structures, not selected by evolution because they are adaptive' (2003: 50-1).

The Person between Moods and Affects

Moods and affects are oppositional extremes in the person's emotional experience and are characterized by different constellations of feelings and by different temporal, intentional, and narrative patterns. In the previous section, I have enlisted the principal phenomenological characteristics of feelings involved in these two kinds of emotional experience. In this final section, I shall put these characteristics of emotional experiences in relation to the concept of a person outlined above. This will be done according to three fundamental aspects of the concept of a person that were also emphasized by Ricoeur: intentionality, temporality and narrative identity.

a) Intentionality

Intentionality is, as we have seen, the aspect of a mental state to be 'of' or 'about' something. The standard phenomenological view on moods and affects is more or less clear on one fundamental difference: moods are unintentional and affects intentional. As Strasser, who has worked out a rich and detailed phenomenological analysis of feelings, writes: 'We must therefore distinguish more carefully than usual between an irritated, angry, happy or anguished mood [Stimmung] on the one hand, and being irritated, angry, happy *about something*, being anxious *over something* on the other. In the first case, irritation is "in the air" without the conscious apprehension of any motive for irritation. In the second case, we are directed, in the mode of feeling irritation, toward something: being-irritated here has an object that is intentionally "meant" […] we shall understand by mood only the felt state of mind [Befindlichkeit], pure being-in-a-mood […] mood must be carefully distinguished from directed feelings' (Strasser 1977: 183/111. Translation slightly modified). This view, however, may be modified by relating the two feeling-states to the person. It is correct to say that an affect as fear is about the particular object of fear (e.g. the bear), and that an anxious mood does not point to any specific intentional object, but manifests itself as an unarticulated background tonality or atmosphere that pervades my whole field of awareness. Nevertheless, my mood seems to affect the way I relate to the world in the sense that it is accompanied by a certain atmosphere in my perceptions. A situation that beforehand would not intimidate me at all now fills me with an irresistible desire to run away and look for protection. The feelings involved in the intentional attitude of my affects are indeed changed by my current mood. My mood is expressed by how perceptions or thoughts affect me. Moods materialize in affects in that I am affected *through* my mood. This may suggest a *covert intentionality* in moods. Whereas affects have a direct and clear intentional object (an object of perception or a thought), moods are characterized by multiple

objects[39]. Whereas affects point to an explicit experience such as a dangerous situation, a happy smile, a beautiful landscape, a difficult task etc., moods, on the contrary, point to my being the person I am in a given situation. Moods can be compared with what Ricoeur calls 'ontological sentiments' in that '[t]hey denote the fundamental feeling […], namely, man's very openness to being' (FM 105/120). We can say that whereas affects point *outward* towards a specific object, moods point *inward* towards my being the person I am. More precisely, one could say that moods contain a *bipolar intentionality* in the sense that they often materialize in a certain affect due to an explicit object, but at the same time point to my being the person I am, and thereby awake questions, doubts, considerations, evaluations, and finally deliberations about my-being-this-person.

Moods and affects interact by means of the person who has the emotional experiences. One way to distinguish a mood from an affect and, perhaps even more importantly, to understand the dialectic between them in relation to the person, is to consider the temporality of feelings.

b) Temporality

The concept of temporality is understood as how the person experiences time and how the existence of the person is inevitably formed and developed in time. Temporality is, therefore, a subjective modality of the concept of time; it is time considered a constitutive part of both the being and the subjective experience of the person. The person changes through time and experiences how the world, other people and herself change in time. Temporality is not time as an exclusively private (solipsistic) or pure cosmological (objective) phenomenon, but both time as experienced and lived by the person and time as working on and with the person.

Moods and affects display different temporal patterns. Affects are often briefer than moods. They captivate me, occupy my whole field of awareness and thereby move me to a determinate action within a restricted period of time. Moods, on the contrary, may last for days, weeks or even years in that they paralyze my thoughts and retain me from acting (sadness) or throw me into weird actions without any thoughts of the past or the future (euphoria). The intensity of the feelings involved in affects demands a concrete action regarding our present situation such as to express our anger, escape the bear, return the happy smile, work on the difficult task, and so on. Obviously, we often do not act out of the affect but retain ourselves from acting out of it. We can dominate the affect by

[39] The psychologist Matthias Siemer is currently working on a similar theory that is called the *dispositional theory of moods* which 'holds that moods are, at least in essential part, *temporary dispositions* to have or to generate particular kinds of cognitions, specifically to make particular kinds of emotion-relevant appraisals' (Siemer 2005: 816-7). He emphasizes that moods are '*multiple-object directed*' (818) and contrasts the idea that moods are 'intrinsically objectless, raw feelings' (842).

a cognitive effort[40]. For example, the irresistible desire to insult or poke a malicious boss may be suppressed by the fear of losing my job. The intensity of the affect then gradually subsides, and I turn my attention on other matters. This, however, does not imply that the affect vanishes altogether. It may remain as a bitter memory that brings forth unpleasant feelings every time it pops up in my mind (Goldie 2000a: 149-51).

The dialectic of moods and affects is complex[41]. Affects may transform themselves into moods and finally become a permanent part of our character; moods may determine affects because they alter the way we are affected by objects and thoughts. Last but not least, a given mood may become an affect when in reflection I articulate it and find its motivations and 'felt causes', i.e., the way it roots *me* in a given *situation*.

An affect can transform itself into a mood that imposes itself on me for days or even longer (grief can transform into a general sadness; anger into dysphoria; boredom into tedium). Thus, a mood may develop out of an affect as the affect itself loses its instantaneous, focused, and motivated character. Also, a mood might not be the product of a single affect and the following action or suppression of action, but a *constellation of feelings* caused by several episodes. Moods (e.g. irritability, sadness, tedium, euphoria) change the way I am affected by the world (and my own thoughts) in that they predispose my field of attention (thus my conscious experience) in a certain way.

And, as we have seen, in the course of time, moods can, in virtue of being dispositional, transform themselves into an inherent and *permanent* part of my self. An affect can develop into a mood, and a mood can develop into a basic emotional tonality. For instance, a dysphoric state can gain such a hold on my person that it turns into a certain trait, for example an irritable, hostile, mean, polemic, misanthropic, or adverse character. This basic emotional tonality is a permanent (often implicit) protention or readiness to (re)act and be affected in a given way, and probably also to develop certain moods more than others. In this way, emotions become an essential part of a person, of one's sense of personal identity. This feeling of sameness comes close to what Ricoeur calls 'character' (se above, pp. 39-42 and 85-87). This basic emotional tonality is usually tacit and I notice it only when it is not there. It is important to notice that all these transformations from affects

[40] This has recently been demonstrated in a neuroscientific study that shows how cognitive reappraisals can affect emotional experience by reducing activity in the amygdala and increase activity in the left lateral prefrontal cortex (Phelps 2004).

[41] I have learned much from on the skilled and detailed phenomenological analyses of affective transitions and alterations made by Stephan Strasser (1977: 203-242/128-60) and Thomas Fuchs (2000: 193-251).

to moods to character occur pre-reflectively and without a deliberate and thematic involvement of the person in the process.

This brings us to the final consideration, namely how the dialectic of moods and affects play a significant role in the process of the person's narrative identity.

c) Narrative Identity

The phenomenological analyses of moods and affects uncover important detail about the basic structures of feelings and the subject who experiences them. However, they need to be complemented by a narrative perspective in order to render information about the personhood of the experiencing subject (Zahavi 2005: 114). As Ricoeur noticed, the subject is characterized by an unrest, non-coincidence, that seeks to recover itself in attestation (or affirmation) of its selfhood. An critical part of this self-attestation is constituted by personhood, i.e., the self as situated in a human society together with other selves (i.e. the humanity of the self) and deeply imbedded in and affected by the historicity of its own being (mediated by the language, traditions, and norms of that society). Personhood involves, as I emphasized above, the radical capacity of the will, that is, that the subject always relates itself to the person that it is; or said otherwise, the subject must will to be the person that it is. Nevertheless, the active, reflective character of human reason (self-evaluation) and volition are intrinsically connected to the passive (pathetic) dimension of the feelings. Strasser puts it like this: 'Human independence is characterized, among other things, by the fact that its feeling-stratum functions as *susceptivum* for all the seeds of reason and, vice-versa, the pathetic powers are ordered to the spiritual realm for their fullness of meaning […] the properly human is the way in which the reason of the person stands to its pathetic powers in a hierarchical relation' (Strasser 1977: 259/174-5). I believe that one way to approach the problematic relation between reason and the person's pathetic powers (emotional experience) is precisely in terms of moods and affects. And narrative identity is an important part of this approach.

Now, Ricoeur's model of narrative identity is complex and works on, at least, four explanatory levels bound together by the notion of temporality. On one level, narrative identity seeks to explore what is particular about human time by individuating a third category configured by both cosmological (objective, anonymous) and phenomenological (subjective, lived) time. On another, it is a tool to approach the hermeneutical nature of the human identity by articulating the historicity of the subject (the subject finds itself embedded in a concrete human world constituted by values and norms shaped by certain traditions and socio-cultural contexts). A third dimension of narrative

identity deals with the intrinsic problems of personal identity (identity and diversity, sameness and selfhood). And finally, it bridges the descriptive and prescriptive level by articulating the normative problems involved in personhood (the axiological dimension or values). I have dealt more closely with these different aspects in part one and showed how the different aspects are closely interconnected. Here I shall only deal explicitly with the two latter ones, personal identity and normativity.

Ricoeur surmises that narratives can articulate the problematic dynamics involved in personal identity (the complex relation and continuous interaction between choices, desire, long time dispositions, habits, character, etc.), because '[w]hat sedimentation has contracted, narration can redeploy' (OSA 122/148). And moreover, he argues that narratives are '*the first laboratory of moral judgment'* (OSA 140/167), since they emphasize the normative aspect of actions and personal identity. These two aspects of narratives explore the fact that personal identity is not only a descriptive process of reidentification, but also a normative problem about how we relate ourselves to the person that we are. However, Ricoeur does not, regrettably, analyze the role of emotional experience in narrative identity. And even though Ricoeur's analyses of personal identity remain the conceptual framework, the following analysis of emotional experience and narrative identity will bear much on the work of de Sousa and Goldie. Although de Sousa does not advocate a narrative approach, he has nevertheless coined a concept that has important implications for narrative theories, namely the concept of *paradigm scenarios*. Goldie, on the contrary, is working intensely on emotions and narratives and has developed, among many others, two helpful concepts that will be employed in the present analysis of moods and affects, *the external perspective* and *emotional resistance*. Further, he gives serious room for the concept of moods in his account of emotional experience, which is in itself a rare and, for my account at least, very precious thing.

De Sousa proposes that we 'are made familiar with the vocabulary of emotions by association with *paradigm scenarios*' (de Sousa 1987: 182). These scenarios can largely be explained as an array of different types of more or less uniform responses to certain emotional objects or emotional events. There exists a multitude of uniform responses that are paradigmatic for human persons. Since early childhood we exist in a world characterized and qualified by emotional experience; for example, we learn to fear glowing hotplates because they might hurt us; or that a genuine smile expresses happiness, compassion or pleasure whereas a frown indicates anger and unpleasant screaming fear or desperation, downcast eyes sadness or grief. Our emotional register is continuously refined and revised in our interaction with the world and other persons (idem: 187); and in an important sense,

'paradigm scenarios indeed define the very character of our emotions' (idem: 183), because emotions are understood in relation to meaningful choices and actions (what Strasser above calls 'fullness of meaning'), and therefore 'their essential role lies in establishing specific patterns of salience relevant to interference […] they are perfectly tailored for the role of arbitrators among reasons' (idem: 200). When we perceive a certain scenario (e.g. a parent hitting her child, a person helping another in need, starving children in the third world, the death of a loved one) we respond emotionally to it, and de Sousa argues that the emotions that the such a scenario elicits (e.g. contempt, compassion, sadness, or sorrow) are defined by paradigmatic scenarios that are again constituted by experience and the biological nature of our being. These scenarios are of course very contextual and furthermore influenced by a multitude of personal and social factors that make the paradigm scenarios deeply individual. Nonetheless, de Sousa argues for an objective account of emotions: 'emotions tell us something about the real world. To be sure, their objectivity is relative to the characteristic inclinations and responses of human and individual nature' (idem: 203). Emotions have a certain rationality that is irreducible to cognitive or strategic kinds, but which is an axiological rationality linked to values and evaluation (idem: 173). Our paradigmatic emotional responses to objects and events reveal something about the values that orient our conduct in the world.

There are at least two important insights to be gained from de Sousa's account. First, the paradigm scenarios allow for an understanding of emotions that include both the biological and personal nature of emotional experience. Individuality and temporal ('biographical') development are kept as essential features of our emotional relation to the world (idem: 100-1) without rendering the biological dimension superfluous; on the contrary, basic affective responses (e.g. fear, anger, panic, lust) 'retain their power even over individuals whose repertoire include the most "refined" emotions' (idem: 184). Secondly, it emphasizes the axiological aspect of emotions. Emotional experience provides information about the world and about me as a person in the sense that they, grossly speaking, reveal what I care about and what I despise. And paradigm scenarios are important for the understanding of our values, because 'our only access to the level of reality that emotions reveal is through the paradigm scenarios that have shaped our world view' (idem: 315).

Now, one way to explore the idea of paradigm scenarios is to dig into the narrative aspect of emotions. De Sousa does not go in this direction himself, although he emphasizes the importance of literature and art for the development and refinement of our emotions (1987: 184; 2005: 351-2). Goldie, however, makes narratives an explicit structure of the experience and explanation of

emotions. His general idea is very similar to de Sousa's picture of emotional experience: 'For each sort of emotional experience there will be a paradigmatic narrative structure […] learning the paradigmatic narrative structure of an emotion, one can come reliably to judge that someone else is experiencing that emotion' (2002a: 105)[42]. Like de Sousa, he emphasizes the personal and developmental aspect of emotions and still maintains that our emotional responses, despite their individuality, often arise from common, paradigmatic structures[43]. Whereas the phenomenological and neurophysiological accounts mainly concentrate upon the *what* and *how* of emotional experience, the narrative structure reveals the *why* (Goldie 2007: 112). For example, why do I have this unpleasant feeling of emptiness when I unlock the door to my apartment, although I know that my family is waiting for me in the kitchen? Or, why do I become angry at my little child even with the most innocent mistake that she makes?

Inspired by the works of the late Richard Wollheim (1984; 2001), Goldie explores the relation between the person and emotional experiences (and actions done out of emotions). To do so, he analyzes what he calls the external perspective or acentrally imagining involved in the narrative structure of emotions (2000a: 196-219; 2002a: 107-8; 2003a; 2003b: 58-9; 2003c: 212; 2005: 135-8). When we, in fiction and in real life,[44] experience the action of another, we immediately connect this action to certain reasons in order to understand the meaning of that specific action. When I see, to use one of Goldie's favourite examples, a person treating a child roughly on the street, yelling and pulling the little one's arm, I immediately experience an *emotional resistance* (2007: 63-7) to that action and to the person doing it. I may try to find all imaginable kinds of reasons that excuse such a cruel treatment of the child (a tuff day at work, an exasperating divorce, the kid's obnoxious character, financial desperation, etc.), but somehow I judge the parent to be a bad person. I may possibly identify myself with her feelings, but I feel a resistance toward her action ('I would never

[42] He is, however, criticizing de Sousa for making the paradigmatic narrative structures constitutive of the different sorts of emotions (Goldie 2000a: 33 note). He says, rightly, that a narrative may involve what is non-paradigmatic, e.g., laugh out of grief or kill out of love (Goldie's examples). I believe though that he misinterprets de Sousa's idea of paradigm scenarios because de Sousa does not exclude that such emotional responses might occur, only that they contrast with what is paradigmatic (normal) of grief or love, namely tears or kisses. The paradigm scenarios or paradigmatic narrative structures are what tells us that an emotional response is appropriate or inappropriate. In this sense, they are constitutive of an emotion. (de Sousa 1987: 185-6).

[43] He writes, for example, 'one cannot grasp the unfolding of a particular emotional sequence without knowing what emotional life is like, and knowing the characteristic ways in which human emotions unfold in response to the vicissitudes of human experience' (Goldie 2003a: 308); or, 'within very broad parameters, there is a rather moving common humanity of emotional responses, but, as with faces, there are subtle and important differences between individual characterizations' (2000a: 109; cf. 2005: 97, 109; 2000a: 85-6).

[44] Goldie argues that 'there is no systematic divide between fictional and nonfictional narratives' (2003b: 66); nevertheless, he stresses that this is not to be taken as a straightforward assimilation of the two cases: 'rather, one should say that a life *can be narrated*, so that the narrative is *about the life*, and thus there remains, in the real life case, but not in the fictional case, the possibility of reference and truth' (2003c: 216).

treat a child like that'). Goldie argues that I do not understand the feelings and actions of another person (e.g., anger, profound grief, charity, loving, or cold indifference) merely by means of, for example, empathy or 'in-his-shoes imagining' (2000a: 194-204; 2005: 137-8), which are both central imaginative processes where I imagine myself as part of the action, but I employ an external perspective on the action and judge according to certain values.

Now, the interesting move in Goldie's account is that he applies this sort of reasoning to a person's own feelings and actions. When I think of my actions (past and future ones) then I engage in a narrative process that reminds of the one we experience when reading literature. One the on hand, I find the *causal* relations between my separate actions (2004: 114; 2007: 111). On the other, I individuate more 'thicker' explanations that often are 'person-specific' (2004: 112) and include an evaluative stance toward actions and events. As I mentioned in part one, there are many of my actions that I do not think about (or even remember), but there are also some that I do think about, and these often reveal something about the kind of person that I am; they involve an evaluative and normative reflection. Those are the ones that I normally understand by means of narrative structures. For example, if I bump into a table so that the coffee cup turns over and ruin my newspaper, it might be combination of bad luck and an instant distraction; however, if I repeatedly bump into things and other people, then this might tell something about my person or, more specifically, my character: clumsy, confused, absorbed in my own thoughts, inconsiderate, and so on. Or a stronger example, if I do not look people in the eyes when I am talking to them, this also reveals some about my person such as arrogance, shyness, or embarrassment. When I reflect upon my experience of the world, the other, and myself, I configure the actions, events, and occurrences according to narrative structures that include both causal and thicker explanations. The concept of person-specific thicker explanations is naturally extremely broad and verges on confusion, but if we (as Ricoeur does) restrict ourselves to the explanations that involve the concept of responsibility and self-esteem, perhaps we might see the relevance of this kind of explanations with regard to emotional experience[45].

How, then, are emotional experience, responsibility, and self-esteem structured in narratives? Following Ricoeur, we may use the concept of personal identity to understand the narrative approach to moods and affects.

[45] Goldie has written a fine, little book that attempts to sort out some of the problems involved in person-specific explanations, where he individuates broad categories of personal characteristics: a) ways of acting, b) habits, c) temperaments, d) emotional dispositions, e) enduring preferences and values, f) skills, talents and abilities, and g) character traits (2004: 11-3).

The temporal aspect of our being is what binds together the various aspects of the narrative approach to personhood. We are changing every second of our life, and yet we remain the same person. Our identity prevails through time, even though our body may alter and deform and our ideas change drastically. Although our identity remains the same, we however change as persons during our life-span. Thus, identity, in the sense of *personal* identity, is not mere sameness. Now, according to Ricoeur, personal identity is formed through a dialectic of two forms of identity: character (idem/sameness) and selfhood (ipseity/keeping one's word). The fundamental trait of character is permanence in time, which Ricoeur, as we have seen, also refers to as character: '[t]he set of distinctive marks which permit the reidentification of a human individual as being the same' (FM 119/144). My character is that in which my feeling of remaining the same in time and through changes is rooted, and that by which other people identify and describe me. A person, however, does not coincide with her character traits; being who she is involves another kind of identity: an identity constituted in selfhood. Whereas my character is formed, at least for some part, involuntarily in the sense that it is determined by my past actions, random events, and contingent factors that are now out of my control, the identity of my selfhood, on the contrary, depends on how I voluntarily relate myself to being a person with this particular character, constituted by a certain past, and situated in a world made up of the anonymous laws of nature, and the societal existence with other persons. My selfhood is constituted by my *active* relation to the person I am (with all the ambiguity that personhood entails); it is I who have the responsibility for my being this person. Here enters the question of responsibility involved in being a person. As Dieter Teichert eloquently puts it: 'Identity as selfhood is linked to a realm where actions are ascribed to agents in the light of ethical norms' (2004: 177-178). It is still me who did that terrible thing in the past, even though it would not cross my mind to do anything like that today. Selfhood entails a kind of self-continuity that implies responsibility not as a contingent, but as an essential component of personhood. This dimension of self-continuity is mainly shaped through narratives.

Here Goldie's work might shed some light on Ricoeur's more abstract conceptual thinking. The idea of an external perspective[46] hinges on the (Augustinian) idea that 'we sometimes need to see ourselves from the outside' (Goldie 2004: 111). Goldie argues continuously for the idea that 'the narrative sense of self' (idem: 117-28) involves a certain detachment from the first-person perspective similar to what we experience when engaged in narratives. We experience both the feelings that the author breathes into the lives of the characters and our own emotional response to

[46] Goldie uses different terms to render roughly the same idea of an external perspective, i.e., *acentrally imagining*, *peripherally imagining* (Goldie 2000: 196; 2003b: 54-59; 2003: 3; 2005: 136-7).

those feelings. When I engage in Raskolnikov's contempt for Alyona, the pawnbroker or Emma Bovary's fascination of a street map of Paris or her fatal attraction to Léon Dupuis, I remain somehow external to these feelings in the sense that I do not identify myself with them although completely absorbed in the feelings themselves. This distance opens up a space of evaluation not identical to a space of reason (Goldie 2007: 110), but somewhat different in that I relate myself to the normative aspect of the feelings. I often experience an emotional resistance to the feelings of the fictional characters, which engenders a normative response to such resistances that reveal our values: 'our emotional responses can reveal to us what we value, and what we value might not be epistemically accessible to us if we did not have such responses' (Goldie 200: 49).

Now, if we apply this narrative structure to our own emotional life, we might come to see 'oneself as another' (Goldie 2003: 312). The fact that what I do and suffer, as my existence in general, is embedded in certain feelings becomes articulated through narratives about myself. We 'loosen' the firm sedimentation of our own character by seeing how the feelings involved in past actions and sufferings have contributed to how we respond to the events of our current daily life. Through our life, interaction with the world and the other has sedimented into a hierarchal organization of what Stasser calls our 'pathetic powers' (feelings) that results in 'a certain dispositional comportment of man' (Strasser 1977: 325/224). Our values are felt in how we respond to what we experience and do. Some of these values are implicit, since they are dispositions and not clearly formulated, but only expressed in the way we engage with the world and other persons. Goldie's central idea is that when we approach ourselves with a narrative structure, we endorse an external perspective on our own thoughts, actions, and feelings, and often we experience the same sort of emotional resistance to our own feelings as to those of other, fictional or real life, persons. Thus, we always evaluate our feelings in accordance with a kind of external perspective that from time to other provoke an emotional resistance to what we actually feel. The fact that we can experience an emotional resistance to our own feelings and doings might suggest that our feelings can reveal something about the fragility of personal identity. We do and say things that we would not do or say. Our self-esteem depends on the responsibility that we have for the person that we are. My ideas, principles, and emotional dispositions are part of what I am, but they can become a burden to me because my personhood is deeply intertwined with the existence of other persons. When I feel in a certain way, this feeling is mine and yet I share it with the persons around me, because there is always something of the feeling in the actions that I do or the things that I say. My feelings are part of my character, even though they often motivate my choices silently and unconsciously (the deceived lover's

complaint: 'I cannot choose to feel what I feel'); in fact, they effectively shape my choices and attitudes. Feelings are at the center of the intersection of the voluntary (activity, spontaneity) and the involuntary (passivity, receptivity), and they need to be unraveled in order to deal with the complex fragility of personhood. The interrelation of self-esteem and responsibility reveals this complexity. Self-esteem draws on my responsibility in relating to others, and yet it also scrutinizes my heart's inmost feelings and desires. My inner life is somehow intimately related to my outer appearance (doings and sayings), because both are part of the person that I am. The narrative approach and the concept of external perspective help engage with the problem of responsibility, self-esteem, and personhood.

There are two important features to be emphasized here. First, how do we understand the notion of external perspective, and, and in particular, what is the relation between such a perspective, personhood, and ethical norms? This will be the central theme of the next chapter. And the second point, which I shall deal with on the concluding pages of this chapter, namely the importance of the relation of moods and affects when dealing with emotional experience and personhood. As I have mentioned, Goldie is among the few who pay close attention to the concept of mood when exploring with emotional experience: 'Moods, too, can affect your thoughts and feelings towards things: moods can crystallize into emotion, becoming directed towards a specific object, and emotion in turn can diffuse into mood' (2002:100).

How, then, do moods and affects figure in the dialectic of character and selfhood developed in the narrative structure of self-experience?

As we have seen in the preceding section, a mood can develop itself into a character trait, i.e., a permanent part of one's sense of personal identity; this transformation often occurs pre-reflectively and without a deliberate and thematic involvement of the person.

However, through narratives, moods can also be incorporated actively, reflectively and thematically into a person's identity. Moods are connected to self-understanding. I understand who I am in the context of my practical engagement, as embedded in a certain world (private as social), and this engagement is primordially enveloped in a certain feelings. My questioning about myself is often elicited by my mood (and by disturbing affects that disclose my mood) before my identity becomes an explicit problem. Moods may disclose to me what word and deeds do not. Feelings are no hindrance to 'cold' rational knowledge, but as Heidegger pointed out our 'understanding' is always attuned, i.e., embedded in a certain feelings. The possibility of self-disclosure, which belongs to moods and affects, is fundamental to self-understanding, because a given mood can point

to a breach in the way I, reflectively, understand myself. I can be locked up in my own way of thinking, chained to my thoughts in such a way that my formulations about myself reflect a wrong or at least problematic understanding of my personhood (Goldie 2004: 125).

Although our capacity to choose to be what (who) we want to be is a constitutive feature of personhood, these choices (evaluations) are always tied to the involuntary aspect of my personhood: "The normative dimension thus pertains to the self that one already is in relating to others, to a world in between, and to oneself" (Grøn 2004: 151).

This complex dialectic between *to be* and *appear to be* (between character, selfhood and world) exposes the fragile, normative nature of personhood: my narrative formulations about myself are constitutive of my person, and yet (as Taylor observes) these formulations can be right or wrong. I can tell a wrong story about myself and therefore live according to this story, but my mood (and its expression in certain affects) may disclose, through its bipolar intentionality, that something is wrong about this story (the external perspective). If we therefore take our emotions seriously and seek to interpret them as disclosing something about ourselves and our situation, then we may gain a tool to approach the problematic notion of personhood.

This, however, presupposes that we understand identity and personhood as essentially anchored in a continuity that demands that we take responsibility for our choices, past as future, and that we do not live our lives as 'episodic creatures' (as proposed by Strawson 1999; 2004)[47]. We cannot base our identity on Lord Shaftesbury's dictum: "The *now*, the *now*. Mind this: in this is all" (quoted in Strawson 2004: 438). In fact, by emphasizing the interconnection between emotions and the normative structure of narrativity we see that feelings may disclose the consequences of living as episodic creatures in the magical now. The feeling of hopelessness emerging when I earnestly try to convince others (and myself) that it was not me who stole the shirt that I am now wearing. It might have been me eons of selves ago, but not the me that I am in this now (or rather 3 seconds ago).

[47] The philosopher Galen Strawson has attacked the notion of narrative identity very harshly. The narrative approach expresses 'an ideal of control and self-awareness in human life that is mistaken and potentially pernicious […] the narrative tendency to look for a story or narrative coherence in one's life is, in general, a gross hindrance to self-understanding: to a just, general, practically real sense, implicit or explicit, of one's nature' (2004: 447). He builds this claim on a thorough analysis of the concept of a self, where he argues for a distinction between self and personhood: 'there are many short-lived and successive selves (if there are selves at all), in the case of ordinary individual human beings' (1999: 100). A self 'is below any cultural variation' (103) and constituted ontologically with a duration up to three seconds, although the phenomenological experience of selfhood might, wrongly, suggest otherwise (111). Selves are characterized are 'SESMETs (Subjects of Experience that are Single MEntal Things)' in the sense that each thought involves a self (118) and 'are physical objects, as real as rabbits and atoms' (120). He the uses this ontology to distinguish between episodic and diachronic creatures, where the diachronics (narrative personalities) tend to be stuck with the past and continuous revising and the episodics (among whom Strawson finds himself) who think that '[t]he business of living is, for many, a completely non-Narrative project' (2004: 448) and that '[t]he 'examined life' is greatly overrated' (1999: 100).

A thematic articulation of my emotions (moods as well as affects) in a narrative structure may help to understand the emotions themselves and thereby approach a fundamental feature of personhood: a person is constituted of both factual (a specific physical constitution, a given world, a certain past, and a particular character) and normative (I can choose to identify myself with the factual or not, but I, somehow, have to relate myself to these facts) features. Emotions open up this dialectic and help us sort it out, if we try to understand (by thematic articulation) the complexity of human affectivity.

It is time to close the chapter on human affectivity. I have spent considerable time on arguing for the irreducible complexity of human feelings, in particular their normative nature, and I will now end by drawing the conclusions of the previous analyses.

Human Affectivity: Feelings and Embodied Normativity

Michael Stocker is one of the philosophers who has spent most time and energy defending the moral importance and complexity of affectivity. He writes that we must understand that 'feelings are complex, intrapersonal and interpersonal, social, and also that they are evaluatively and morally important […] feelings are not mere feelings; there is little that is mere about them; and this explains how what is mistakenly thought of as mere feelings can be of vital evaluative and moral importance' (Stocker 1996: 54). My account is deeply inspired by his seminal work about the irreducible affective dimension of care, concern, and interest. The principal difference lies in my attempt to articulate the interrelation between moods and affect within the emotional experience itself, and further to establish the outline of an argument for the importance of such an articulation in relation to personhood. The following chapter will extend the argument of the evaluative and moral importance of emotions by developing what Goldie has called the external perspective in the narrative structure of emotions. I will attempt to link such a perspective to what I have called the practical space of reason in order to argue for the indispensable relation between subjectivity and ethics. I shall argue that subjectivity becomes, to some extend, an empty concept when considered as a feelingless (minimal) sense of agency and ownership; and further, if the subject is considered as a feeling-loaded agent, questions of personhood and ethical dispositions becomes indispensable for a comprehensive account of subjectivity. My argument is basically that the concept of subjectivity is a thoroughly normative concept. But before turning to that central chapter, I will, in way of conclusion, articulate the most important insights arrived at in this long chapter.

I will first recapitulate, and thereby clarify, my account of human affectivity. I structure my analyses on one of Ricoeur's most basic ideas (inspired as he is by Kant), namely that subjective experience is constituted by a transcendental synthesis of reason and sensibility, that is, subjective nature is constituted both by spontaneity and receptivity. On the one hand, the subject is capable to think and act spontaneously without being induced by anything but itself. On the other, the subject is affected by that which is other (the external world, the other subject and even itself) through the senses and, of particular interest in the present context, the feelings. Subjective experience is neither pure reason (spontaneity or activity) nor pure sensibility (receptivity or passivity), but a synthesis of the two. Furthermore, I agree with Ricoeur's other basic idea that actual subjective experience is always characterized by affectivity. The subject always experiences the world and itself as loaded with feelings; there is no such a thing as a feelingless experience or feelingless reason. This is because human experience is, at its inmost core, a synthesis of spontaneity and receptivity. In this general sense, affectivity accounts for the receptive aspect of being human. We are affected by the external world, time, and, in particular, by the other persons (through our senses), but also by the physical state of our body, our thoughts, ideas, principles, memory, and so on. In a more specific sense, however, affectivity accounts for what we know as emotional experience or emotions. As mentioned briefly above, I use emotions as an umbrella term for the variety of affective phenomena (feelings, cognitions, affects, or moods) that we call emotional experience. Emotional experience involves two fundamental elements, cognitions and feelings, and emotional theories have been fighting over which element is the dominant one. I give them equal weight in the sense that, depending on the emotion, the cognitive and feeling aspects vary and differ as to which element is the salient feature of the emotional experience. Some emotions are characterized as more prominently cognitive whereas others are closer to mere bodily feelings. The decisive argument is that all emotions depend on both cognitive and feeling elements; we never experience a pure cognitive nor a pure feeling emotion due to the fact that our experience is primordially a *synthesis* of spontaneity and receptivity. We always contribute something of ourselves to the emotion, our worldview, our desires, and ideas; and still the emotion captures us, renders us passive, and imposes a certain feeling upon us that we cannot escape. The emotion tell us something about the world (the informative claim or the objectivity of the emotion)[48] and about our own relation to that world, the

[48] 'The information I have in mind is not confirmed to the world outside the subject: if I am angry at you, my anger tells me not merely that you have done something wrong but also that I am in a position to assess and react to it. The many kinds of information implicated make for a rich mess, but one mess, after all, can still make one topic' (de Sousa 1987: 114)

other, and ourselves (the subjectivity of emotions)[49]. My proposal is then that to access and understand the complex variety of human emotions, we must first phenomenologically individuate the constellation of feelings involved in a given emotional experience. The tool to do so is a model which classifies the emotion within one unified scale from affects to moods according to the phenomenology (intentionality and temporality) of the feelings involved, and then we must contextualize and personalize the feeling in a narrative structure that reveals the cognitive elements in that emotional experience and the complex ambiguity of the factual and normative aspect of personhood (what the emotion means to me, and what it reveal about the world, the other, and my personhood). The feeling aspect is a necessary, but not sufficient part of the explanation of emotions (Strasser 1977: 194/120). This phenomenological-narrative approach puts my account close to Goldie, and yet my emphasis on the affect-mood model and the normative-factual ambiguity of personhood makes my account quite different from his.

Second, why the emphasis of moods and affects as extreme poles of human emotional experience? I have made this model because I believe that by gathering our emotional experience within a unified scale of feelings and cognitions that is held together by the extreme poles of moods and affects we get an effective tool to phenomenologically analyze and categorize the different emotional experiences according to their intentionality and temporal duration. This categorization helps the narrative analysis of the emotion, because we come to terms with the object(s), phenomenological manifestation and temporal duration involved in the emotions. Furthermore, it makes the developmental aspect of emotions clearer than it is possible if we distinguish strictly between moods and emotions as if there were talk about two completely different phenomena. In addition, it evidences the salience of cognitive elements involved in the emotional experience. Moods often seem to lack a distinct object (I cannot put my finger on why I am sad) but which might suggest a more complex cognitive element (a covert-bipolar intentionality characterized by being multi-object directed), whereas affects often involve a clear object (I fear the bear), brief temporal duration, and more simple cognitive elements. Finally, the emphasis on moods are important when it comes to the values involved in or revealed by emotions. The multi-object directed nature of moods constitutes a more fine-grained approach to emotions and values, because whereas more simple-layered emotions (e.g. affects such as joy, anger, surprise, fear, disgust, or

[49] 'I and the world are embedded in an undivided experience of totality. Disposition is the feeling of I-and-world together' (Strasser 1977: 188/115); or 'Now, one's own I is one pole of concrete individuality which is first brought into relief in I-feeling, self-feeling, feeling of particularity, and so forth. On the other hand, the world is a massive, concrete and individual object to which the I is necessarily related in its struggle for self-realization. This, too, is first revealed in feeling' (idem: 192/119).

panic) reveal more direct and clear values (e.g. offence, safety, well-being, survival), moods (e.g. anxiety, melancholy, emptiness, indifference) reveal my complex relation to the values that I hold, to the person that I am, and to other people. Whereas affects reveal my values, moods may effectuate questions about my values, and the way I relate myself to the world, other, and my own person.

This brings me to the third and final conclusion about the embodied and normative aspect of emotional experience. As we saw above, the neurophysiological dimension of emotions emphasizes the importance of emotions in person's comportment in the world. Damasio advocated the somatic-marker hypothesis that emotions highlight certain possible choices in the process of decision-making and thereby restrict our field of attention that would otherwise be immense and impossible to comprehend. Without emotions, our experience of the world would be flat, and everything would seem indifferent. I agreed to the general hypothesis but refuted the implications of the theory, because it reduces the notion of subjectivity to a simplistic notion and the aspect of emotional normativity to a question of evolutionary benefits, primary emotions and physiological changes. This, however, does not reduce the importance of the neurophysiological dimension. Panksepp showed how emotions might be conceived as basic affective programs (raw feelings) that determine primary instinctual behaviors and reveal basic ecological values embedded in human life; further, he emphasized that these values affect all our cognitive skills in the sense that human feelings are immersed in these basic affect programs[50]. A clarification of such programs might help understand the cognitive impenetrable nature of many of our emotions. Our experience of the world is always embedded in deep affective structures that point to both our general mammalian nature and our specific human nature and concerns.

This tells us something about the normativity involved in emotions. Emotions reveal the complexity of values. For example, physical values appertaining to the physiological constitution of my being (vital values); personal values that reflect my own (sometimes) peculiar values, social values (etiquette and aesthetic values) that influence how I comport myself in society with other persons, and finally ethical values that reflect the practices and norms by means of which I act toward to other people. The question of ethical values will be the topic of the next chapter which hopefully will make it clearer what I mean by the fundamental relation between subjectivity and ethics. In this long chapter, I have tried to show the complexity of human affectivity, and how our feelings may reveal something about the values by which we lead our life. Feelings show us what

[50] Actually, Goldie proposes that we substitute the talk about basic emotion with talk of affect programs (Goldie 2000: 106).

we care about, and how our personhood is a normative question initiated by how we feel about ourselves and the world. One might say that emotions are primary indicators of the normative dimension of human life; our feelings reveal an embodied normativity that complicates the rigid distinction of fact and value, because they emphasize that our human experience of facts in and about our existence is always loaded with values.

Chapter Two

Ethical Experience and Ontology

In this concluding chapter, I shall tie together the foregoing analyses and answer the central question of the work, namely about the relation between subjectivity and ethics. I have spent a long time on human affectivity because I believe that it discloses a basic normativity that is too often disregarded in both theories of subjectivity and theories of ethics. But subjectivity is obviously much more than revealed by the affective dimension. Here I shall focus on why ethical experience is of primary importance to a theory of subjectivity, and further how such an experience might reveal something essential about of human ontology. I pick up where Ricoeur concluded, namely with ethical experience and an ontology of care. I have extended considerably the analysis of the affective dimension in comparison with his analyses, because I consider the absence of a development of the notion of affectivity to be one of the weak points in his theory of subjectivity. Although the following analyses continue to draw heavily on his groundwork, I think that an elaboration the idea of the practical space of reason will clarify notions and arguments that remain somewhat inarticulate and obscure in his extensive writings. The idea of an ontology of care is only mentioned peripherally, and without a clear argument for such an ontology his otherwise strong arguments about the primacy of ethical experience remain a curiosity in the contemporary debate of subjectivity and ethics. Further, I develop and use his basic ideas and analyses in a different context than where they are commonly employed.

I draw heavily on the work of two other contemporary philosophers to articulate my conclusions. The first is the Simon Blackburn whose work on practical reasoning has helped me understand that a naturalization of values does not necessarily involve a reductive account of neither subjectivity nor normativity. The other is the Charles Taylor whose extensive writings on the primacy of morality, in particular the idea of 'the self in moral space' has inspired me to the development of the notion of a practical space of reason.

But before turning to the principal analyses, it might be a good idea to take a brief look back to two great philosophers whose work has produced many of the concepts and arguments that are used in the contemporary debate about subjectivity and ethics. Such a move backwards can help to clarify, position, and refine the following analyses.

A Preamble: Kant and Hume on Emotions, Ethics and Subjectivity

Emotions, ethics, and subjectivity have been intertwined in western philosophical tradition throughout more than two millennia. As with so many other philosophical issues, Hume and Kant are still the ones to consult for historical and argumentative insights into contemporary debates. They are notoriously famous for their completely opposite views on the relation between emotions and subjectivity. On the one hand, we have the view expressed in Hume's famous words: 'Reason is, and ought only to be the slave of the passions' (Hume 1978: 415) and further '[m]oral distinction cannot be derived from reason' (idem: 458). This psychology of action is the foundation upon which Hume builds his ethical theory. It is not that reason does not have anything to say; on the contrary, there cannot be any moral without reason, but it is not reason to have the first or the last word on ethical conduct. Reason alone does not move us to act. Morality, therefore, is grounded in a universal sentiment of humanity due to 'the original fabric and formation of the human mind' (Hume 1975: 172). In Hume we find an appraisal of emotions as the source of morality (and of action in general) in human nature. The reason why we care about the welfare of the other person is because of a universal feeling of sympathy, 'the humanity of one man is the humanity of every one, and the same object touches this passion in all human creatures' (idem: 273). This is the principle upon which all moral reasoning must be founded.

On the other hand, Kant looks suspiciously on the emotions as a part of our sensible nature and therefore not fitted for morality. All feelings die out after a while, whereas reason and its concepts stand forever (Kant 1998b: 129-30/157-8). True morality is grounded in reason, and is therefore not subject to the whims of the subject or the pressure from a sick society. Reason alone can guarantee a moral conduct, since only by following the rules of reason are we, as he writes in *The Metaphysics of Morals*, able to dominate our feeble animalistic nature and become more human or, perhaps even more importantly, become worthy for the humanity that harbors in our (rational) nature (Kant 1996: 151/387). This hidden humanity in human nature is reason, and the guiding rules of conduct are expressed in the categorical imperatives of the moral law which 'is given, as it were, a fact of pure reason of which we are a priori conscious and which is apodictically certain though it be granted that no exact example of it can be found in experience' (Kant 1998b: 41/47). Obviously, Kant is well aware that '[n]o human being is entirely without moral feeling, for were he completely lacking in receptivity to it he would be morally dead; and if (to speak in medical terms) the moral vital forces could no longer excite this feeling, then humanity would dissolve (by chemical laws, as it were) into mere animality and be mixed irretrievably with the mass of other natural beings' (Kant

1996: 160/401). However, this natural and innate gift must be cultivated and educated by the firm principles of reason otherwise they very quickly turn ugly and end up as pathological. So Kant does not at all ignore feelings and emotions in his ethics, which one also notice when reading his anthropological and political works that are closely related to his ethics. He simply does not count on them (Kant 1998a: 17/404-5). Emotions are way too futile and instable to be considered as a part of a serious ethical philosophy. Therefore, they cannot be a motive in moral psychology. One day my conduct may be motivated by the 'moral feeling' in my nature and the next day, just as naturally, I may feel it best to think more of my own welfare and less of that of the others. Without the categorical imperative and the rules of pure (practical) reason my conduct is formed and controlled by 'the slings and arrows' of heteronomous forces such as feelings, imagination, contingent events and facts. Feeling are part of the natural world and, therefore, they cannot be a motivational factor for a rational being that completes its autonomy in acting in accordance with laws issued by a non-natural reason with which all rational creatures are endowed (Kant 1998b: 59-61/69-70), and by means of which they can aspire to become a member and perhaps even a sovereign of the 'Kingdom of Ends' (Kant 1998a: 41/433).

The connection between imagination and feeling is important since it is employed with equal force in both Kant and Hume, although with drastically different outcome. Whereas Kant (in his ethical writing) mainly points to imagination and feeling in relation to the pathological phenomenon of lust and other egoistic tendencies (the right use of imagination is guided by reason), Hume sees feelings and imagination as fundamental for the concept of sympathy and true (i.e. not selfish) interest in other people. He writes in the *Treatise*: 'We blame equally a bad action, which we read of in history, with one perform'd in our neighbourhood t'other day: The meaning of which is, that we know from reflexion, that the former action wou'd excite as strong sentiments of disapprobation as the latter, were it placed in the same position' (Hume 1978: 584). Without imagination we could never reach what Hume calls 'the common point of view' (Hume 1975: 272) or 'the natural sentiment of approbation and blame' (Hume 1978: 579); that is, the position from which we look upon our feelings and sentiments without any direct relation to ourselves.

On the other hand, however, Kant still recommends travels, living in a 'cosmopolitan' city like Königsberg, reading novels, history, biographies, and going to the theatre as ways of gaining important insights into the nature of human beings. Therefore, in some way, he cherishes and approves of the power of the cooperation between imagination and feeling (Kant 2006: 4-5/120-2). But it is different matter in ethics where such a combination cannot be allowed to direct our

conduct. Here reason stands alone as the sole defender of our humanity. Not aided by imagination or feelings, or by a combination of the two.

What kind of picture do these two metaethical positions give of emotions? Well, on the one hand, Hume presents a theory that tries to integrate emotions as a fundamental part of ethics. The nature of human subjectivity plays an essential role in what we judge as a 'good' or a 'bad' action. Hence, the ontological status of moral value is subjective, i.e., dependent on the particular constitution of human nature. We come to know about values by paying due attention to emotions and to how these emotions express themselves in the social conduct of human beings (the epistemological aspect). Reason alone can never discern a value from a fact. Thus, reason could never motivate us to do anything if it were not for emotions. On the other hand, Kant strictly separate ethics and emotions. Human subjectivity does not affect the ontological status of moral value. Moral value is objective, not dependent on human nature (which is constituted by sensible features such as feelings and other bodily sensations). The good and the bad action is not perceived by our eyes or our feelings, but known by reason alone in abstraction from anything empirical. They could be understood by a brain in a vat or by inhabitants from an emotionless Twin Earth, were these only endowed with reason. Therefore, emotions cannot, or at least should not, motivate our conduct in interacting with other persons. This has interesting implications for the relation between subjectivity and ethics.

In Hume and Kant we have two different approaches to the question of subjectivity and ethics. In their theoretical works, both philosophers express a quite skeptical view on the question of subjectivity. Hume's famous words, 'I never can catch myself at any time without a perception, and never can observe any thing but the perception itself' (Hume 1978: 252), tell a story about the missing self that is always something else (a perception 'of heat or cold, light or shade, love or hatred, pain or pleasure' (idem)). Kant's view is nothing but clear. The question about the self was the reason for the rewriting of the *Transcendental Deduction* and the *Paralogisms of Reason* in the second edition of the *Critique of Pure Reason*, and Heidegger wrote a book on this problem, which he thought to be the fundamental question in Kant (1990). Kant seems to evade the question and yet believe it to be the most important. What is the self? It is the common root (Wurzel) of sensibility and reason, the hidden power of the soul, the logical unity of apperception, and the metaphors goes on, but in the second edition he concludes that 'I as intelligence and thinking subject cognize my self as an object that is thought, insofar as I am also given to myself in intuition, only, like other phenomena, not as I am for the understanding but rather as I appear to myself' (Kant 1997: 259/155). We only have an indirect knowledge of subjectivity in both philosophers, and the subject

remains fragile and obscure when we try to define it by a theoretical approach. It is in their practical philosophy that we find their full-blown theories of subjectivity. In Hume the subject becomes a human subject by means of human sentiments and feelings, and in Kant the subject becomes a subject 'in itself' (*an sich*), a person, by listening to the clear voice of reason that declares that a human subject can never be treated as a means, but only as an end in itself, a person What makes a subject a human subject is, in both philosophers, the moral dimension of subjectivity.

Nevertheless, their respective ethics have different implications for a theory of subjectivity. Kant develops an ethics that reacts against what he believes to be the sensible flaws and inconsistencies of the human subject, whereas Hume grounds his ethics in that same sensibility. Kant sees an irremediable gap between sensibility and reason that demands an ethics strictly separated from the theoretical investigations into the nature of the human subject. The normative is of a non-natural nature, and therefore it depends upon practical and under no circumstances on theoretical reason (Kant 1998b: 100-2/119-21)[51]. Hume, by contrast, claims that the normative dimension cannot be approached without a thorough exploration of the natural nature of the subject. In fact, it depends on that natural nature. This is not to say that the normative is the natural (Hume was the first to point to the naturalistic fallacy (Hume 1978: 290)), but as Jonathan Jacobs says: 'We can explain the source and basis of moral value and judgement by reference to naturalistic features of human beings' (2002: 127). Feelings and desires, being an aspect of our sensible nature, are therefore either abandoned as pathological or cherished as fundamental in their ethics. Where does this put us with regard to the relation between ethics and theories of subjectivity? Kant puts a watershed between the two whereas Hume thinks them inseparable, considering emotions essential to both human subjectivity and ethics.

In contemporary philosophical debate, we might see ethics as somewhat alien to a descriptive theory of subjectivity; theoretical and practical analyses are done on different levels by philosophers who are specialized in one area but not in the other. What do normative features have to do with a factual description of the human subject? This might seem a fair objection to an objective, non-

[51] It is worth to notice that Kant's ethics can be approached from another perspective than through his explicitly ethical publications. His political and anthropological writings (accessible both in published works, manuscripts, and the so-called reflections) provide a much more refined view on the nature of emotions and ethical reasoning. A good example is his long defense of sensibility in the *Anthropology* (2006: 34-54/143-67). Peter Strawson points to the same fact in an argument where he gives Kant's theory of subjectivity more credit than that of Hume exactly because Kant has shown, with the analyses of the transcendental unity of apperception and the paralogisms, that subject-identity needs no criteria, neither empirical nor rational, to actually exist. Subject-identity is a logical presupposition for experience, and despite his notorious 'neglect of the empirical concept of a subject of experience […] The point is, however, that nothing in Kant's account excludes, and everything in it invites, such a supplementation' (Strawson 1966: 169-70; se also Strawson 1959: 102-4). I return to some of Strawson's arguments in chapter two.

natural ethics as the Kantian presented above, but would seem less obvious with regard to a Humean ethics. And this is due to the nature of emotions. In the following section, I will show how subjective experience is primarily ethical experience and that this is due to the particular structure of our experience as a practical space of reason.

The Practical Space of Reason: Values, Concerns, and the Other

The phenomenological analyses of affectivity showed that the subject is always in some kind of feeling-state that reveals the being-in-the-world of the subject, i.e., existence is always enveloped in feelings. We cannot escape the feelings when analyzing the nature of subjective experience. From a neurophysiological perspective, Pankseep emphasizes the same primacy of the feelings by pointing to the fact that although we are '[d]eeply cultural creatures with a cognitive apparatus, arising from vast neo-cortical territories that can associate perceptions and ideas with feelings, that remains unmatched in the living world […] those abilities would disappear, like a fading dream, without the solid platform that arises from subcortical abilities we still share with other animals' (2004: 58). According to Panksepp, the basic, non-specific human, affective aspect of our being is fundamental for our more elaborated parts of our nature such as aesthetics, culture, or ethics. The philosopher Dan Zahavi points to the same basic affective feature in subjectivity while discussing why autistic persons have serious problems in intersubjective relations: '[w]hat we are confronted with here is not a lack of a theory of mind, but a lack of an immediate, pre-reflective, or implicit understanding of emotional expressions and the unwritten rules of social interaction' (2005: 221). Our being a subject cannot be separated from emotions and desires, neither with regards to selfhood nor to intersubjectivity.

It is exactly with the question of intersubjectivity that subjectivity and ethics become related. But often this relation first becomes evident at a higher (i.e. less basic) level of intersubjectivity, where we reflectively form theories of other minds and then try to find out which kind of conduct we should choose and which we should decline. But an ethics that takes the emotions seriously would see it otherwise. It argues for the pervasiveness of the normative. Normativity is present at every level of selfhood, even in the minimal self-awareness: '*there is no getting behind ethics*. It comes unbidden. It comes with living' (Blackburn 1998: 2). We act in accordance with norms although these norms be can more or less explicit, more or less articulated. The recent neurophysiological discovery of mirror neurons in higher primates (monkeys and humans) seems to empirically support or at least enforce the theory of pervasiveness of the normative. Mirror neurons are a particular

class of visuomotor neurons originally discovered in area F5 of the premotor cortex in monkeys. Later, they were also individuated in humans and are now believed to exist in other brain areas as well. What make them special is that they are activated by both proper action and observation of the action of others. I understand the action of another before cognitively relating myself to his or hers action[52]. Our relation to others is determined by a pre-reflective and automatic understanding of their action and their concerns and needs. According to a hypothesis called 'the shared manifold hypothesis', we understand our world as a 'multidimensional, 'we-centric' shared space' (Gallese 2003: 172) that secures, by means of an embodied simulation at the neuronal level, an immediate recognition of other as persons with intentions and feelings like ourselves. One of the pioneers in this research, the Italian neurophysiologist Vittorio Gallese, writes: 'Our capacity to conceive of the acting bodies of others as *persons* like us depends on the constitution of a shared meaningful interpersonal space […] Embodied simulation constitutes a crucial functional mechanism in social cognition, and it can be neurobiologically characterized' (2007: 9).

The discovery of mirror neurons is interesting, since it promises to break with the otherwise solipsistic tendency in contemporary cognitive neuroscience to study the anatomy and neurophysiology of the brain in isolation from the interaction with other brains (Gallese 2003: 172; 2006: 15-6). Our world is already a world of concerns and needs, ours and the others. There appears to exist an 'unmediated resonance' among individuals that secures an emotional understanding among persons (Goldman and Sripada 2004: 207-8). Without exaggerating the importance of the discovery, one can at least observe that the pervasiveness of the normative does not contrast with current research. One of Ricoeur's central points is that most, if not all, actions involve and influence the presence of other persons (pp. 90). Normativity is simply instantiated by the fact that our existence is deeply interrelated with the existence of other persons; if we take into account the basic affective values that Panksepp argued for, we might say that this is so even in the deep subneocortical layers of the brain,. The most basic structures of living organisms are characterized by normativity in form of primary feelings of pain, anger, care, seeking, and pleasure. This is, of course, an extremely more complex process in humans, due to, among other things, the ontological fragility of our nature (the vacillation between to be and appear to be); however, there seems to be a

[52] A fine and accessible introduction to mirror neurons is found in (Rizzolatti and Sinigaglia 2007), although the presentation is somewhat scared by a serious lack of reservation and perspective on the circumstances and result of the discovery. A more modest presentation of the discovery would be to characterize it as a 'promising heuristic tool' for future empirical investigation, as it is done elsewhere (Rizzolatti, Fogassi, and Gallese 2001: 669). The speculative effort surronding the research in mirror neurons often outweights the exactness and relevance of the empirical data. I thank Dorothée Legrand for pointing this out to me (cf. Legrand 2007: 470-71).

phylogenic continuity among mammals that involves the capacity of embodied simulations at a neuronal level which procure the experience of the world the fundamental feature of being a 'we-centric shared space' (Gallese and Umiltà 2006: 28). The affective dimension, our feelings, or sentiments as Hume would call them, reveals this normative feature of experience.

Thus, if we accept (as I do) that emotional experience reveals many of the concerns and needs of human subjects, then it is the job of ethical reflection to understand, structure, and evaluate these in relation to choice and action. Simon Blackburn, deeply inspired by Hume's sentimentalism, has developed an ethical theory from this way of looking emotions and ethics. His view is known as a version of expressivism (Ayer, Hare, Gibbard). The theory argues that moral judgements are expressions rather than assertions of facts or beliefs. Moral judgements express feelings and attitudes and do not depend on the realist's obsession with truth-conditionals (Blackburn 1984: 191-196)[53]. The theory follows Hume a long way, but put an emphasis on language that were not present in his ethics (Blackburn came to ethics via philosophy of language). I will not go into detail here, but mainly sketch how this ethical theory depends largely upon a theory of emotion, and vice versa, how a theory of emotion must take into account the moral dimension of emotions outlined in this theory. This interrelation tells us something about the reality of human experience that will help us reinforce the notion of a practical space of reason. But in order to show this, we will need to look at some of Blackburn's basic arguments.

Blackburn's Quasi-realism and the Experience of Values

Blackburn insists that emotions reveal a large amount of the subject's concerns and desires and explains how some of these are considered as values and others not. Our concerns and desires are ordered hierarchically: 'We should think in terms of a staircase of practical and emotional ascent [...] The staircase gives us a scale between pure preference, on the one hand, and attitudes with all the flavour of ethical commitment, on the other' (1998: 9). A value differs from a pure preference or other kinds of desire in that '[t]o hold a value is to have a relatively fixed attitude to some aspects of things, an attitude with which one identifies in the sense of being set to resist change, or set to feel pain when concerns are not met' (1998: 68). Values are, as in Hume, neither natural facts nor supernatural entities, but something that the subject feels towards the world, itself and other

[53] In ethical discourse '[t]he existence of the verdict, of course, depends on the existence of those capable of making it, the existence of the truth depends on nothing (externally), and on those features that make it wrong (internally) […] there is no doctrine to express relating the truth of the verdict to the existence of us, of our sentiments, or of rival sentiments' (Blackburn 1988: 1978).

subjects. This is not to be misunderstood as if values were to depend on emotions only. This is not so. Reason has a lot to say in the structuring and education of values and preferences. The theory only insists that without the emotional aspect there would not be values in the first place. Emotions qualify the natural facts as preferences, concerns and values[54]. Blackburn calls this kind of doing ethics practical reasoning in the sense that ethics cannot be something external (by a reference to a natural or non-natural realm of values) to how the subject normally thinks, feels, and acts. Values are part of how we think in our practical relation to the world, ourselves, and other subjects. And in this picture, emotions play a prominent role because they reveal our concerns and values. That is why a theory of emotions is important to this kind of ethics. The more we understand about emotional experience (in particular the interrelation of activity and passivity, cognition and feelings, moods and affects), the better we are able to structure our own desires and concerns in accordance to those of others.

Blackburn grounds his ethical analysis in what he calls quasi-realism which he positions in between realism and anti-realism in ethical debate. Realism (e.g. Kant, Plato, Moore, Nagel, Korsgaard) is the position that understands qualities such as goodness and badness or rightness and wrongness as being actual properties in the world, or at least committed to some kind of truth. Said otherwise, our belief that something is good and another thing bad must reflect some actual state of affairs in the world. Our beliefs about the normative qualities of things are dependent on truth-conditionals. For example, if Louise is a kind person, then she must possess the true qualities of kindness. But where do I find these qualities that make up her kindness? The realist believes that we can find them independently of what we think or feel, that is, independently of our subjective nature. They need not be of some obscure metaphysical order, but 'rather just the truth about what we and others should do and want' (Nagel 1986: 139). However, the problem for the realist is how to account for such a truth of values and conduct. If we are not to search for it in God's authoritative commandments, a platonic ideal order, or in the Kingdom of Ends, how is it possible to argue for the true existence of ethical qualities and propositions?

There is a common agreement among (most) realists that we cannot find the truth of moral values and conduct in the natural order as it is revealed by the empirical sciences. Therefore we cannot understand moral values in empirical terms, but we would still have to argue for the objective

[54] He writes, for example:'[The] talk of something being good, or something being a reason for action, is a kind of reflection of a motivational state of mind: the fact of something weighing with you. This motivational state of mind is not a simple belief. It is not a representation of some aspect of the world. It is a *reaction* to representation of the fact of the matter. It does not itself pick out some fact of the matter. Hence it is not strictly speaking a state of the mind that is either true or false, any more than a desire for coffee is either true or false' (Blackburn 1999: 283).

existence of these same values; otherwise we slide into a subjective position which means trusting the fragile nature of our feelings and attitudes. One way to avoid the subjective treat is to choose a cognitive approach as we saw it in Kant. Reason can tell us what is wrong and what is right. This has been expressed as 'a claim to law or universability' so that 'I need to will universally in order to see my action as something which *I do* […] the function of the normative principles of the will, in particular, is to bring integrity and therefore unity – and therefore, really existence – to the acting self' (Korsgaard 1996: 228-9). The normative principles are sovereign because, just like Kant's categorical imperatives, they govern and hold down self-interest, disdain, selfishness, pleasing sentiments, and other bad feelings and attitudes. In establishing firm principles like this, we secure that our actions are guided by disinterested rules of conduct and not by the whims and flaws of our fragile human nature. If we want to know if Louise actually possesses the true qualities of kindness, we have to see if her actions are done out of principles and not just feelings of kindness. The cognitive realist finds the objectivity of moral values (Louise's kindness) in the universal principles that are epistemically accessible by the work of reason alone.

This might seem convincing if we exclude the question of naturalism (and that of moral psychology[55]). Suppose we are naturalists (as I am) and therefore want to argue that the empirical sciences have something important to learn us about ourselves, our nature, and the nature of the world in which we live our life, how do we then find 'room for ethics, or of placing ethics within the disenchanted, non-ethical order which we inhabit, and of which we are part' (Blackburn 1998: 49)? One option is to be a moral anti-realist (e.g. Ayer, Hare, and Mackie). Perhaps the most famous anti-realist is the moral skeptic John L. Mackie. He bluntly states that '[i]f there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe' (1977: 38).

Whereas the realist is committed to some kind of objectivity, the anti-realist needs not to, and can therefore freely opt for subjectivism, relativism, or even skepticism in moral discourse. We cannot lock our commitments onto any objective standard, because even though we normally think that our claim of rightness or wrongness possesses some kind of objective validity, this is shown, on closer scrutiny, not to be the case. Mackie calls this sort of denial 'an error theory' (idem: 35), because it demonstrates that we are in error every time we use moral language as if it referred to some real, objective standards, rules, or principles. The problem about anti-realism is that it often leads to

[55] Remember Ricoeur's scepticism regarding Kant's feeling- and desireless account of motivation which he calls 'ethical dualism' (pp. 37; FM 79/95). He writes bluntly that 'no motive can incline me if it odes not impress my sensibility (pp.28; FN 122/117)

relativism. If there are no real, objective standards, structures, or conditions in moral discourse, then the slip to 'anything goes' is very easy. How are we to determine if 'shmoralising' is worse or better than 'moralising' in moral disagreements if we have no standards to judge by (Blackburn 1984: 186). Therefore, Blackburn characterizes moral relativism in the following harsh manner: 'The moral relativist believes that arguing about ethical truth is like arguing about the true location of the rainbow' (1998: 298). The fact remains, however, that whereas we seldom argue about the location of rainbows, we very often argue about what is the right thing to do. Moral discourse is an attempt to solve moral problems, dilemmas, and conflicts; and although our beliefs in the standards by which we judge something to be wrong or right might be shown to be in error, 'it does not follow that the error infects the practice of moralizing, nor the concepts used in ways defined by that practice' (1985: 150). This needs explanation. And that is exactly what Blackburn provides with his argument for a third position that he has named quasi-realism. Basically, the central argument goes like this: in order to solve moral dispute or disagreement, we need to argue and practice *as though* there were actual moral truth to be had, because our values and norms are partly dependent upon our own sentiments, feelings, and attitudes, and partly on the facts that we find in the world. Our values are expressions of our attitudes to the world, ourselves, and the other; in other words, we must 'understand the ethical proposition as a focus for practical thought' (1998: 77; cf. 1980: 29-30; 1984: 209; 1988: 167).

This may seem, in the eyes of some, a rather weak attempt to squeeze ethics into a natural order while still respecting the peculiar nature of our moral beliefs (McDowell 1997: 159-66). Nevertheless, I believe that Blackburn's account of quasi-realism actually centers on an important feature of our understanding of the world, namely that our experience is constituted by both the unquestionable existence of the world and our peculiar subjective nature. We might remember that this was one of the basic structures in Ricoeur's theory of subjectivity. The world does not depend on our way of taking it; it imposes itself on us through our body and our being situated in it. And still, our interaction with the world (and the other subjects) is also determined by how that world appears to us by means of our five senses, conceptual reasoning, and our feelings. We owe much of what we know about the world (that it is not flat) and ourselves (that we are not the center of the universe, and that smoking might eventually kill you) to the empirical sciences; therefore, these are not to be considered as a obstacle in ethics (or as reasons for ethical skepticism as in Mackie's error theory); on the contrary, they can be a help to further understanding the nature of both the world that we live in and our own nature. What is important, however, is to understand that our values and

norms are still different from the other properties discovered by the empirical sciences, since they express what is peculiar about human experience of the world. This does not mean that they are independent of the factual world, because, as Blackburn insists, 'moral properties must be given an intelligible connection to the natural ones upon which they somehow depend' (1984: 187). But there is no use in running around looking for truth-conditionals for our conduct and values, since we must leave 'ethical properties and propositions alone with their own specific identity. They are counters in our transactions with our values, just as a piece of money is a counter in financial transactions. To understand the value of a piece of money it is no use starring at it. It is necessary to understand the processes of human economic behavior […] We need the same eye for whole processes of human action and interaction. We need synthesis, not analysis' (1998:50).

Now, quasi-realism is very close to projectivism in the sense it surmises values and ethical qualities to be a product our attitudes to the world, our responses to things and people around us; we project our 'inner' feelings, ideas, concerns onto the external world, and thereby the non-ethical world becomes the world that we are concerned about; it becomes our world. This, however, does not exclude that the world at the same time can have its own laws that our feelings and attitudes cannot change. On the contrary, our ethical dispositions are exactly molded by this causality. For example, we do not blame a person losing his memory because he is becoming old or has suffered brain hemorrhage, or having his hat blown of on a windy autumn day. But then again, we do blame him for pushing and pulling himself forward on a crowded street in January because he has forgot to put on warm boots so now his feet are freezing, and he wants get to indoors as soon as possible. The quasi-realist insists that there are facts which are independent of our projections or expressions of attitudes, but he also states that some facts cannot be explained only by these facts or by the anonymous laws governing them because 'the word 'fact' also has an uncorrupted English sense: it is the fact that there are colours, and there are many facts about them' (Blackburn 1987: 57). What the quasi-realist opposes are not the empirical sciences, or the validity of their theories, it is the realist's reduction of reality to an external world untouched by human attitudes, feelings, and perceptual capacities; a world in which colors and taste are characterized as 'secondary' qualities with a lesser degree of existence than the firm substance of mass and form. It is even worse with moral feelings such as shame or guilt or ethical properties such as goodness and badness, since these are very hard to fit into a 'square world' of primary forms and masses. It was this problem that led Mackie to become a moral skeptic and Kant to invent the 'Kingdom of Ends' where ethical values are indeed primary and the 'unnatural' pure reason the ticket to come along.

Ethical properties, feelings, and values are as real as gravity and supersonic airplanes; indeed, sometimes they even seem more real or at least more actual in our lives, because they effectively shape the way we think about and act in the world; they concern me more frequently than gravity and supersonic airplanes. And experiences of ethical values such as good, bad, guilt, or shame are different from experiences of mere preferences and idiosyncratic propensities by the fact that they necessarily involve the presence of other people; they must weight more on my thoughts and action than my secret little whims, because it is by means of such strong values that I orient myself in my daily life. If I pay too much attention to my own private feelings and concerns than to those of others then I am, rightfully, exposed to other people's reprobation. This, in turn, normally leads to feelings of guilt and shame, which is a good and natural thing, because '[w]ithout these emotions, the motivation to act well is diminished' (Blackburn 1998: 20). It is in this sense that our feelings reveal what is important to us, what we care about. I may care about my freezing feet in a cold January afternoon, but, hopefully, I also care about what other people think of me, therefore I know that I should not push and pull to go faster, and if I have actually done so, guilt and shame are natural consequences of such a behavior. To exclude feelings and attitudes from ethics is what Stocker, with his usual perspicuity, has called 'moral schizophrenia', because '[n]ot to value what moves one also bespeaks a malady of the spirit' (Stocker 1976: 454). It is schizophrenic because of the split between motive and reason. My feeling of kindness or respect cannot be a motive for action, because it does not fulfill the criteria of a real moral reason. Both utilitarianism (Mill) and deontology (Kant) is guilty of such a charge, since '[w]hat is lacking in these theories is simply – or not so simply – the person. For, love, friendship, affection, fellow feeling, and community all require that the other person be an essential part of what is valued' (idem: 459).

Quasi-realism takes the dispositions, attitudes, and feelings of the subject so seriously that it inflates the status of our evaluative predicates so that our projections onto the external world are treated *as if* they were facts like gravity, causality, and freezing feet. In fact, Blackburn adopts a minimalist conception of truth for his quasi-realism that 'allow us to end up saying ´It is true that kindness is good´. For this means no more than that kindness is good, an attitude that we may properly want to express' (1998: 79). The fact that our feelings, attitudes, and actions can be true and good does not entail an analysis of truth-conditionals, but an eye for the workings of our practical reasoning. How do we experience and act in a world that contains both anonymous events, which cannot be ascribed moral qualities, and moral actions that in fact are characterized by such ascriptions? In a quasi-realist picture, there is no tension between a naturalistic world description

and a moral discourse; they both tell us something about the facts of a world that is full of many heterogeneous things such anonymous causality, sexual desires, unconscious inherited dispositions to alcoholism or violence in form of DNA structures, social oppression, happy faces, bad luck, and, last but not least, malicious and kind feelings. But we must not forget that even though values and ethical qualities depends on and are in relation to the facts of the empirical sciences they have their own special kind of identity, their own nature. In this Moore was right, all appropriate ethical reasoning depends on how well we understand that ethical notions are what they are and not something else. They are different from other notions that we are aware of (Moore 1903: 17). In fact, the difficult task is to retain the central and often unsurpassable insight of the realist's accounts about values and ethical properties without sliding into mystical and inappropriate explanations of ideal orders (Platon), parochial virtues and small poleis (Aristotle), infinite Kingdoms of Ends (Kant), solid natural utility (Mill), or an intuitionist capacity (Moore)? Blackburn proposes that we consider ethical propositions about things and persons to be constituted by the fact that they reflect our concerns; that is, our values, feelings, dispositions, and attitudes that comes with living (1998: 80). To understand the meaning of these concerns, we need understand how we think, feel, and act with respect to the concerns that characterize the persons that we are. And here Blackburn introduces a notion that comes very close to the notion of the practical space of reason, which I proposed in my reformulation of Ricoeur's theory of subjectivity, namely *practical reasoning*. He does so, because he insists that '[e]thics remains essentially practical, a matter of attitude, disposition, and emotion' (idem: 65).

In the concluding sections, I will pursue this idea of practical reasoning, but no longer with particular attention to Blackburn's analyses, since I am here primarily concerned with the relation between subjectivity and ethics and not with nature of ethical reasoning itself. I will use his quasi-realism together with Goldie's external perspective on emotional experience in order to give a final articulation of, and hopefully strengthen, my central argument that subjective experience is primordially configured as a practical space of reason.

Experience as Structured by the Value of the Other's Concerns

My reformulation of Ricoeur's theory of subjectivity developed a central idea that I named the practical space of reason. Even though Ricoeur never used or discussed a similar concept, I believe that his theory implicates that subjective experience is primarily structured according to a hierarchy of heterogeneous values organized by the practical use of reason. The subject experiences the world

as a meaningful organization of things and events that matters to the subject as a person. Things and events have a value for the subject in the sense that the subject is always, somehow, concerned with what it experiences. Experience is never without feeling, and the subject's sense of agency and ownership, 'the ipseity of the experiential dimension' or 'a minimal sense of self' (Zahavi 2005: 146), is always permeated by normativity. One might say that to be a subject is not only to act and experience that it is I who act and experience, but also to feel that this sense of self matters in some way or the other. I always *relate* myself to the world, the other, and my being this person, although this is often in a pre-reflective, unmediated sense of the word. Ricoeur individuated this basic relatedness in the nature of the subject as an 'affective fragility' in the heart of the subject. Subjectivity is fragile in nature and characterized by unrest, often even conflict, between vital and spiritual desires, between the originating affirmation and the existential difference that characterize every subject. This conflict is, first of all, revealed by the affective dimension, that is, the feelings that characterize the being of the subject. Ricoeur, however, is not very clear or detailed about the affective dimension, although he continues to emphasize it fundamental importance. Therefore, I have spent considerable time on developing a more elaborate account of affectivity, both from a philosophical and neurophysiological perspective, in order to sustain the importance of feelings in a theory of human subjectivity. Feelings reveal values and the pervasive nature of normativity. Panksepp's account of affective consciousness insists that feelings are constitutive of basic ecological values in form of instinctual affective behavior that influences the subject both in its reflective and pre-reflective relation to the world, the other, and itself. De Sousa, Solomon, Strasser, and Goldie claimed that feelings have an informational importance that surpasses mere subjective whims and inclinations. I used their accounts to develop a model which includes moods among the feelings that reveal the normative character of personhood. I believe that the notion of personhood is important in order to fully understand the normative significance of more diffuse and atmospheric feelings involved in moods. Emotional experience (that is expressed by the concept of emotions) contains both clear affects and diffuse moods, and we need a model that explains both in order to understand what we mean by affective experience.

In this chapter so far, I have tried to show how emotions are important for the ethical discussion of value and conduct basically because, as Blackburn has show, emotions show us what we *actually* care about, our concerns about and attitudes to the world and other persons. Ethical reflection is a practical matter about how we handle our emotions and organize our concern in our daily interaction with other persons. Blackburn's quasi-realism insists on the importance of the subjective

nature of our experience. We should not consider our affective experience as less informative or less factual than our 'cool', detached reasoning. Both have their hold on human nature and both are part of our experience of a world qualified by values. The realist's sharp distinction between subjectivity and objectivity in talking about values crumples in front of the experiential dimension of human concerns. Subjective experience, because it is emotional experience, always influence the way we talk about value and ethical dispositions. Our actions are done out of reasons, and reasons are constituted by different motives, and, as Ricoeur, Blackburn, and Stocker point out, we cannot disregard emotions as concrete motive for action without ending up in what Stocker calls 'moral schizophrenia'. In fact, I doubt that any motive can be completely feelingless. As Hume pointed out, emotions are part of 'the original fabric and formation of the human mind'. Values and the normative dimension of subjectivity are as real and factual as any aspect of the subject. The values and normative outlook by which we lead our life may be more difficult to handle and decide on than a broken arm or the boiling temperature of water, but they exist in the same factual manner as such things. It has recently been argued against the metaphysical refutation of the reality of the self (cf. Strawson 1999 or Metzinger 2003) that 'why not rather insists that the self is real if it has experiential reality and that the validity of our account of the self is to be measured by its ability to be faithful to experience, by its ability to capture and articulate (invariant) experiential structures' (Gallagher and Zahavi 2008: 213). The same argument holds for normative experience, since, according to my arguments, the self is in the most basic experiential structures an affective and normative being. This is so because subjective experience is primarily structured as a space of practical concerns that are indeed organized and handled by reason, but pre-reflectively experienced as values that affect us through our feelings. These concerns are hierarchically organized in terms of the heterogeneous values that we experience in our being in the world; or, as Blackburn calls it, 'a staircase of practical and emotional ascent'. When I see a pencil in front of me, I do not experience it as small object with specific physical characteristics, some primary (form and extension) and some secondary qualities (color and taste), but as a thing with specific affordances that enable me to perform certain actions[56]. Things or events are always experienced in some kind of relation to

[56] This has been analyzed thoroughly by Heidegger as 'zuhandenheit' and 'vorhandenheit' (Heidegger 1996: xx/66-76). He writes, for example: 'Again, the surrounding world makes itself known. What appears in this way is not itself one thing [Zuhandenes] among others and certainly not something *objectively present* [*Vorhandenes*] which lies at the basis of the useful thing at hand [zuhandene Zeug]. It is "there" before anyone has observed or ascertained it […] That the world does not "consist" of what is at hand can be seen from the fact (among others) that when the world appears in the modes of taking care which we have just interpreted, what is at hand becomes deprived of its worldliness so that it appears as something merely objective [Nur-vorhandensein zum Vorschein kommt] […] When the world does not make

myself that is revealed by the feelings that the thing or event awakes in me. I value what I experience before I reflectively relate myself to that thing. The values that I experience may not be very clear or articulate, but they affect me on a pre-reflective level. It is not possible to separate clearly between the objectivity and subjectivity of the value, since the experience of values is constituted, as we have seen, by my peculiar subjective, personal nature *and* the objective, anonymous nature of the world. But to say that a value is only real if it can be detached from the nature of the subjective experience that constitutes it is absurd. Therefore, Blackburn developed the notion of quasi-realism to save both the reality of values and the importance of subjective experience as a constitutive element of values. If we want to talk about the reality of values as independent of subjective experience, then we should at least concede that there are objective structures in subjective experience that reveal more than just the particular values of individual subjects. The philosopher Lars Christiansen describes beautifully this complex relation between objectivity and subjectivity of experience: 'Although experience in a way can be said to come from ourselves, it is not self-imposed in any disqualifying manner; we are not the masters of experience, it is in this regard as objective as the thing itself. One could just as well criticize daylight for falsifying the things, because it does not come from them but from the sun'; or even more acutely some pages later, 'Here music can teach us something. A scientific ideologist who overhears someone praising the benevolent sun would refer to the fact that the sun is simply a glowing bubble of gasses, and how could anyone praise gasses? In the same manner, one could imagine an unmusical person rejecting a violin concert with reference to the fact that here it is actually just a matter of rubbing horsehair against the bowels of a cat' (Christiansen 1980: 71, 74. My translation).

Experience configured as a practical space of reason means that our experiences are always affected by how we value the experienced objects and events. We perceive objects and events as meaningful things and events that somehow concern us and reveal the values by which we orient ourselves in the world that we live in. Our concerns are expressed in our feelings, dispositions, and attitudes. And our values are expressed in our concerns, in the sense that we 'can call the things that matter to us our *concerns*' (Blackburn 1998: 123). When we talk about reality, we should not forget that *our way* of experiencing is also a reality that we immediately relate ourselves to, although some might find it a lesser reality than the one disclosed by the empirical sciences.

Nonetheless, the distinction between subjectivity and objectivity is still essential for understanding the nature of our values and concerns. But we must understand what we mean by subjective and

---

itself known, that is the condition for the possibility of what is at hand not emerging from its inconspicuousness' (70/75).

objective. To say that my concerns reflect only what *I* care about, not regarding what others care about is completely to misunderstand the nature of subjective experience. One of the points that Ricoeur returns to again and again is that '[*t*]*o say self is not to say myself*' (OSA 180/212). Subjective experience does not only reflect the narrow preferences and predilections of the individual subject; it is constituted by the world and especially the other subjects. The practical space of reason is therefore not confined to the point of view of the individual subject. The experience itself is of course individual in the sense that I am the one who has the experience, but what is important is the experiential nature of what I experience. Subjective experience is structured as a common experiential space in which I am immediately aware of the presence of other subjects. My perspective is not the only perspective; and I know this not by means of reflection, but it is a pre-reflective awareness that is developed by social exchanges in early infancy (about two months old) and becomes an essential feature of human experience of the world. We experience the world and our actions in it from both our first person perspective and a '*third person or allocentric perspective*' that together are 'combined to form new triadic meanings' (Rochat 2007: 16). I am aware that the chair has a back that I cannot see but is visible to another subject in different position, or, in the same manner, I am pre-reflectively aware that my action has consequences for other persons because my experience and action is situated and embedded in space constituted by the world, the other, and myself. Hence, the practical space of reason is noninferentially intersubjective, because perception, intention, and actions are structured in relation to alterity: 'the three dimensions "self," "others," and "world" belong together, they reciprocially illuminate one another, and can be fully understood only in their interconnection' (Zahavi 2005: 176). The intersubjective structure of subjective experience is significant, because it tell us something about the objective nature of some of our concerns and the quasi-realistic nature of values

We met this idea earlier in Goldie's analyses of emotional resistance and the external perspective. When we do an action we might feel an emotional resistance that cannot be explained by other than our capacity to immediately see that action from an allocentric or external perspective. This reveals the complex importance of feelings in our experience of the world. My feelings are not only an expression of my private concerns, but also the expression of the concerns of others. To talk about concerns involve the other subject's values as well as my own. The external perspective indicates this complexity in the concerns that structure my experience. My values are developed and refined by living in a world that does not obey my will (I do become older no matter what I do, and therefore I perhaps learn to cherish the value of things that seemed ridiculous to me when I was

younger, for example sitting in a quiet cornfield counting crows). In the same way, social interaction with other people affects the way I am concerned with what goes on in my life. These concerns may be implicit and cognitively impenetrable, but are nevertheless revealed in my attitudes and feelings (my moods and affects). The practical space of reason is constituted by heterogeneous values in the sense that my concerns for things and events are not restricted to only *my* perspective but always affected by alterity. This alternity is an essential part of the constitution of selfhood (involuntariness, passivity, body, time, world, and the other, as we have seen in Ricoeur) and is expressed in the structure of our experience. The external perspective is an inherent feature of our experience, because we are pre-reflectively aware of otherness when we perceive, act, feel, and think. Goldie characterized it (inspired by Wollheim) as acentrally imagining to emphasize that the experience is not egocentrically structured but involves a perspective different from our own. The fact that we experience a phenomenon such as emotional resistance indicates that this external perspective does not come after the immediate experience as part of a later reflection upon what happened. We can experience an emotional resistance to our own actions, feelings, and thoughts. Such feelings often arise before we reflectively relate ourselves to what has happened. It is layered deep in our own way of experiencing and acting in a world shared with other persons. Of course, this immediate reaction to things and events, actions and sufferings, thoughts and memories needs a reflective approach in order to be fully understood, but this is not the point here. What is essential is that the embodied and situated nature of our experience is structured in such a way that, whether we like it or not, our concerns are inherently constituted by values that transcends our own point of view, our own likes and dislikes. This explains why Ricoeur called value a quasi-concept (pp. 95). We cannot determine the nature of a value a priori or universally, as Korsgaard does she when 'tries to move *from* the fact that we have reflective distance from our impulses *to* the requirement that we conceive our reasons as universal, at least in the formal sense' (Korsgaard 1996: 225). Human values are not formal. They depend on the subjective nature of human beings, social and cultural contexts (in China they actually eat man's best friend), historical development (thunder and lightning is no longer considered wrath of the gods, but only bad luck if one happens to be outdoors). However, this is not the same to say that value is subjective in a relativistic sense like 'this is right because it is right for me, and you cannot tell me otherwise or convince me that I am wrong!' Our values are grounded in the concerns of other. When I experience the world, I pre-reflectively see things and events from different perspectives, and my own concerns are therefore necessarily interwoven with those of other persons. Thus, our feelings, dispositions, and attitudes

(i.e. our character) are an essential source of normativity. They reveal how ethical questions are impossible to get behind in understanding the nature of subjectivity. We are concerned about the other, whether we like it or not. On the other hand, subjectivity is a necessary part of an approach to ethics, because values cannot be understood without taking into account the subjective (quasi-conceptual or quasi-realistic) constitution of values. To understand why we act the way we do, we have to understand the nature of our being human subjects. Blackburn says it much better than me: 'It is only when we have human nature under some control that human ethical nature comes under control' (1998: v).

This may seem a somewhat thin argument for the relation between subjectivity and nature. Therefore, in the concluding section, I will try to strengthen the argument by grounding it in a development of what Ricoeur called an ontology of care. This will also be of help in the last part where I discuss the specific problems of neuroethics, moral relativism, the conception of nature.

An Ontology of Care: Personhood and the Conflictual Nature of Subjectivity

The ethical dimension of subjectivity seems to demand the use of 'thick' ethical concepts such as generosity, courage, good-sense, prudence, kindness, cruelty, duty, moderation, and so on. Such thick concepts describe the character of a person, the dispositions, and ultimately its ethical status as a moral agent. Some think that to reduce thick concepts to thin concepts, such as good, bad, right, true, wrong, is to dissolve ethical qualities into what they are not. We cannot explain why a person is generous by refereeing to simple minimal facts. We need to see the thick ethical concepts as basic irreducible concepts to orient ourselves 'in moral space' (Taylor 1989: 28). To know what to do when faced with moral problem, I have to know the basic thick concepts of moral discourse. These concepts cannot by explained by facts that are different. Cruel is cruel and kind is kind, and these concepts are firm reference points no matter what we do. They are principles that can not be diluted into more thin qualities such as good and bad. An ethical minimalist sees it differently. He or she sees the thickness as 'overdramatic […] It implies too ready a tendency to diagnose the discussion and re-evaluation that are essential of ethical activity as exercise in talking past each other' (Blackburn 1992: 299). The fact is that most of the time we do not talk past each other. We know what we talk about even though we never use the words kind or cruel; that is, we do not need a special concept to determine the quality of an action. Our ethical disposition (values, rules of conduct, and firm principles) depends on our attitudes and feelings, on what we care about, and not on our ability to understand the semantic properties of lexical expressions. We must look at the

practical aspect of human life and the basic structure of care in order to understand ethical dispositions. In the previous chapter, I argued that a concept of personhood is essential for understanding human affectivity, since feeling always involve the nature of the beings that experience them. And human subjects are persons. However, Ricoeur noticed that this fact involves the task of becoming a person. The subject cares about being a person, and very often a certain kind of person. Feelings direct and focus our attention to things that matter to us as a person, because 'having emotions and noticing and attending are closely interrelated' (Stocker 1996: 85). In this way, our feelings and attitudes reveal this problematic fragility in the concept of personhood. One cannot fail to be a person, but very often we fail at being the person that we want to be.

The central problem is how to explain the values that we live our life by. Are they grounded in a special moral ontology that governs human life and the concept of personhood, or are they simply expressions of our feelings, attitudes and dispositions, and therefore part of a natural world that knows nothing of kindness and justice, gold littered clouds or the value of Braque's cubistic trees? I have argued that subjective experience is constituted by kindness, laughter, and gold littered clouds as well as by indifferent forces of physics and evolution, because our experience is configured as a practical space of reason structured according to heterogeneous values derived from both the world, the other and the self. Our cruelty and goodness are just as natural as the falling of raindrops or diarrhea. I do not think that we need a special moral ontology to thicken our behavior in order to find generosity, benevolence, brutality, or evil. All we have to do is to respect the complexity of our feelings and refine the ontology that explains our being in the world. However, before I explain my idea of such a refined ontology, which includes both the fragility of personhood and natural values, I will take a brief look on one of the most influential and acute defenders of an explicit moral ontology, namely Charles Taylor.

The Moral Space and Strong Evaluations

First of all, I must start by saying that I am inspired by Taylor's work on subjectivity, personhood, and morality. The only problem that I have with his writings is a small detail that can, however, have important consequences. This is his view on naturalism. In over forty years, Taylor has written against naturalistic explanations of human behavior, and I am in perfect agreement with his rejection of reductionist accounts of agency and personhood. But naturalism is not a synonym for reductionism. To be a naturalist is not necessarily to hold that all can be explained by the mechanisms of physics and chemistry, but that these branches of the empirical sciences can help us

understand the nature of human beings. Taylor is well aware of this and would not disagree with an ambition to find a convergence between, for example, neurophysiology and human subjectivity (1971: 178-9). The problem is that he makes such a convergence quite difficult. He holds a strong realist idea of morality and an even stronger aversion to anti-realist or quasi-realist ideas about the projective nature of ethical value in a disenchanted physical universe (1989: 59, 78). He maintains that we cannot explain moral experience by other than real moral values that have nothing to do with the naturalistic inspired 'gut reactions […]similar to those of baboons' (idem: 15; cf. 6) or 'weak evaluations' (1976: 116). Real moral values are expressed not by our feelings, but by strong evaluations that secures a 'subject who strongly evaluates desires, goes deeper, because he characterizes his motivation at greater depth. To characterize one desire or inclination as worthier, or nobler, or more integrated, etc. than others is to speak of it in terms of the kind of quality of life which it expresses and sustains' (1977: 25). Any attempt to naturalize or dilute these strong evaluations in a quasi-realist picture is to undermine morality (1989: 6).

I believe, on the contrary, that we can maintain ethical commitment and the quality of human life without necessarily adopting the strong vocabulary that Taylor argues for. Although it may seem a contradiction, I surmise that his analyses can be used in a lower, i.e., less antagonistic, tone to argue for the naturalistic nature of values. His perceptive insights into the relation between personhood, emotions, and values can be used to argue for a practical space of reason instead of a moral space. The notion of a moral space structured by strong evaluations and thick concepts put more weight on the reflective nature of human values than these can bear. The values by with we live our life are too heterogeneous and discordant to be enveloped in consciously endorsed moral principles. Many of our less noble and more selfish concerns affect our thoughts, feelings, and actions in a much stronger way than is allowed in a moral space. It is not just a matter of purifying my person for such non-moral concerns. My identity actually depends on those more 'dirty' evaluations just as much as on the more strong qualitative (clearly moral) ones. As Ricoeur observed, our existence is organized by practices that are basically more ethical than moral; that is, the idea of living well is not primarily a moral one, but about having a good life. The characterization of 'good' is not only a matter of moral norms and obligations, but being able to live according to the values that include such moral norms as a *part* of my own concerns. My hierarchy of values must of course reflect the norms that govern my interaction with other persons, but they are intrinsically intermingled with my own ideas and nasty little desires that may be in direct contrast to the norms, which I hold to be part of the person that I am. My existence is more than moral behavior, and often moral behavior is

more a wish than a fact because, as Blackburn writes, 'Jago is just as natural as Mother Theresa, and on a head-count perhaps more so' (1998: 48). To be a naturalist is to rely on the empirical sciences for important help in understanding human behavior, feeling, and thought, but it is also the temperate observation of ourselves and others in the daily routines of human life. Here (alas!) we seldom find actions done out of consciously endorsed moral principles but out of a heterogeneous variety of motives and values (from the most depraved aggression to the most warmhearted unselfish deed).

There are several reasons why I have chosen to analyze subjectivity by means of a practical space of reason instead of 'the self in moral space'. As mentioned above, I think that there are many non-moral sources of the self. Further, the reformulation of Ricoeur pointed in that direction. For despite all his emphasis on moral constitution of the subjectivity, Ricoeur has keen eye for the self(ish) in the concept of a person, the tension between self (the I) and selfhood (the person). This is the non-coincidence in the heart of the subject that makes personhood fragile and vulnerable. To be a person is a task of becoming a person; the subject has to balance the heterogeneous values that orient its existence in order to live as a person together, with, and for other persons[57]. The practical space explains such heterogeneous values and the means-to-end reasoning that characterizes so many of our concerns. And, at the same time, it fulfills the formal criteria of Taylor's moral space. Taylor argues that we need a moral space to explain our (modern) moral thinking, which has three principle axes: 1) sense of respect of others, 2) a full life, and 3) dignity (attitudinal) (1989: 15). These three axes make up the three dimension of the moral space, which is the 'background, explicit or implicit, for our moral judgements, intuitions or reactions (idem: 26). This space is constitutive of human agency because 'stepping outside these limits would be tantamount to stepping outside what we would recognize as integral, that is, undamaged human personhood' (idem: 27). The difference between Taylor's space and mine is that I do not think that we have to include strong evaluations as a basic constitutional feature of what explains human agency or personhood. This is not to be understood as if we could exist without strong evaluations. On the contrary, they are a central part of what it is to be a human person. I only say that the practical space of reason can account for such evaluations in term of a more 'weak' explanation, i.e., in terms of

[57] This is the central point in his criticism of Lévinas, namely that an ontology of subjectivity cannot be grounded in an asymmetrical conception of self and other. Therefore, he writes that 'Self-imputation […] is now inscribed within a asymmetrical dialogic structure whose origin lies outside me' which means that 'No middle ground, no between, is secured to lessen the utter dissymmetry between the Same and the Other' (OSA 336/388, 338/390). This dissymmetry results in a 'hyperbole of separation' that 'renders unthinkable […] the distinction between self and I, and the formation of selfhood defined by its openness and its capacity for discovery' (idem 339/391).

concern or care and practices. I found this idea on the affective dimension of personhood. Our feelings are too heterogeneous and pervasive to be held within the armor of morality. Even our most conscious endorsed principle is rooted in feelings, although this might not be the clear intentional emotions that we usually associate with emotional experience (fear as well as benevolence). Feelings have many expressions (consciously and unconsciously), which I have tried to show by my extended analysis of moods and affects. Our identity hinges on our feelings, because the complexity of *personal* identity is, as we have seen, often cognitively impenetrable. My identity is not only a matter of how I consciously relate myself to the person that I am. Sometimes feelings reveal what is hidden beneath my actions, emotional reactions, dispositions, or character. I cannot understand what is wrong with me, and that that absence of understanding is a constitutive part of my identity as a person.

Further, as Taylor rightly observes, feelings are what secures the significance of the things for me (Taylor 1981: 107-9). In fact, Taylor's analyses of the emotional aspect of significance, which he calls 'the import of a given situation' (1977a: 49), is critical for my own account about the irreducibly informative nature of emotional experience. Taylor calls this the 'full-blooded activity' that makes a human being and a frog different from a calculator, because 'things *matter* to living beings' (Taylor 1988: 450). Our feelings reveal what we care about, and we do care about more than just strong evaluations. We care about getting the kid to school, smoking our cigarettes, our (lack of) hair, the personal life of Britney Spears, the size of our television, our father's bad health, the balance of our bank account, and so on. These are all part of what makes us the persons that we are. Naturally, dignity, fullness of life, and respect of others are more important than a new beautifully expensive sweater, and therefore these strong issues should instantiate the supreme compass by which we orient ourselves in life. Nevertheless, the identity crisis that Taylor describes as 'an acute form of disorientation, which people often express in terms of not knowing who they are, but which can also be seen as a radical uncertainty of where they stand' (1989: 27) is not only the lack of framework of 'stable significance', which is in itself a rare phenomenon, but also due to the structural fragility of subjectivity. I believe Taylor is completely right when he emphasizes the moral aspect of the fragility of modern identity, but I do not think that he is right when he argues that a naturalist humanism is defective in the sense that it hinders universal benevolence because it does not identify it with the Christian agapē (idem: 516-8). If we want to understand the normative constitution of human subjectivity, we cannot approach it as a *moral* question about orientation in terms of strong evaluations. Normativity is present in every action, feeling, and thought of the

subject, but we need a more deflated form of morality in order to see how morality is something that we cannot leave out of a factual description of the human subject. I do not doubt that strong evaluations are some of the most firm and reliable supports for our fragile nature, and that they cannot be left out of any serious account of subjectivity. Nonetheless, I am convinced that the practical space of reason can account for such evaluations without excluding the naturalistic understanding of values. Subjective experience is configured by a hierarchy of heterogeneous values that comes with living in a world, which does not change because I want it to, and together with other human subjects that demand that I recognize their personhood. But these values themselves are not clear and neither is my hierarchy, because, somehow, that world and other people are still what I take them to be. My experience of the world is always constituted by this difficult relation between my self and alterity (body, world, and other). I sometimes (too often) value things wrongly, and this is revealed because my experience and my own being a person is constituted by that which is not me (the world and other persons). My experiential and interactive relation to the world is primarily practical in the sense that my existence depends on how I interact with the world and other people on practical level. Thus, I have argued that subjective experience is primarily configured as a practical space of reason according to heterogeneous values expressed in my concerns or rather in what I care about.

Now, my final argument will be about the ontological significance of personhood in relation to the notion of care and the conflictual nature of human subjectivity. Hopefully, these concluding pages will reinforce my argument for the irreducible relation between subjectivity and ethics.

Ontology and Personhood: The Conflictual Nature of Subjectivity

The notion of a practical space of reason has been the central theme up until now and the notion that binds the different analyses together. Subjective experience is primarily configured as a practical space of reason that is structured by the heterogeneity of values generated and shaped by the world, the other, and the self. Human experience always implies an experience of values revealed by what we care about, our concerns. Experience and action are so interlaced that experience of objects and events is basically practical. We perceive objects and events as *things* and *situations* that concern us; in short, we care about what we experience.

The affective dimension is a primary source of information, since our feelings make the world, the other, and our existence in general matter to us. Experience without feelings is unthinkable, because, as Stocker puts it, ‘human life is through-and-through affective and in large and important

ways constituted or perfected by affectivity. Further, human life, in its various forms, is judged in terms of affectivity – importantly, by whether and how it succeeds or fails in developing, expressing, and satisfying affective needs' (Stocker 1996: 325). There is no getting behind our feelings if we want to understand human subjectivity. In fact, affectivity is an important part of the basic structure of subjective experience (receptivity), and as such it influences all our capacities (perception, reason, will, and action). To separate reason and feeling is to misunderstand the nature of human experience, since human experience is always emotional experience. Further, it is an expression of mistrust to how we actually think about the world and ourselves, and a reduction of the heterogeneity of feelings to 'mere bodily' induced phenomena. All experience is embodied experience, which means that we reason in terms of our bodily constitution, so excluding mere feelings is to want to get rid of a fundamental part of our nature. There is no such thing as a mere bodily feeling without meaning or relevance to our existence. It may be just as relevant to care about involuntary sweating palms as to respect my college's coffee cup. Feelings can, however, be more or less appropriate and even more or less right according to the situation that I find myself in. Say that I laugh when I see a serious accident just because it was ironic that the car hit a lady who were adjusting her hair and caring about her looks; 'now she will never have a bad hair day again'. That is not just an inappropriate reaction, it is a wrong reaction. It goes against all that human society stands for and cherishes. The reply 'But you cannot neglect the fact that it was ironic and a damn funny thing, if we, for a moment, suppose that nobody got hurt' will not change the fact that it was a wrong reaction *because* a person got hurt. Fortunately, persons normally do not find such a tragic event funny, which tells us something not trivial about the nature of human care. Care expresses the nature of subjectivity; to imagine a subject who does not care about anything is difficult, if not impossible. Ricoeur aligned this care the originating affirmation or attestation (the will to live, the will to say yes), but he also stressed that subjective care is never confined to our own concerns, because we live together with other persons who care about things as we do. Feelings express our care about…, and to understand the heterogeneous nature of feelings (from moods to affects) is to understand how care holds together our experience and action in the world; it binds together the being of the world and our being persons who care about that world. Care is the ontological ground from which we understand what it is like to be a human subject. Our experience and existence is practical in the sense that we experience, feel, think, and act according to what we care about. Our care expresses the values by which we live our lives, just as feelings express our care. Feelings, care, and values are inseparably interrelated in subjective experience, which is why I

have stressed the affective dimension so much in my analyses. Emotional experience is the starting point in understanding both the experiential and ontological nature of subjectivity. Reason and rationality is fundamental in our relation to the world, the other, and to ourselves, because we understand, develop, and refine what we care about in terms of makes sense and what does not. But reason works in and with affectivity like the sculptor works with the amorphous stone. There would be no reason to reason if we did not care about reasoning; and we care about reasoning, because we want to know better what we care about and what to do with things, events, and persons that we care about. Reason and rationality are fundamental to our life, nobody denies that, but it is important to understand that 'the strict and categorical contrast between rationality and emotionality is misleading and tendentious' (Stocker 1996: 104).

To see how care expresses the ontological condition of subjectivity, we need to bring in the concept of personhood. First of all, the notion of care covers both the biological and the psychological dimension of subjectivity. We care about our basic biological needs just as we care about love, kindness, beauty, and friendship. We cannot separate our personhood from the biological nature of our being. This was one of Panksepp's central points. Basic affective instincts are present even in our most refined spiritual thought (it has always been considered a sacrifice to become a monk or join a hunger strike; in fact, that is exactly what such choices want to demonstrate). However, as argued repeatedly, affectivity is not confined to such instinctual needs. Feelings are also intentional. They do not only characterize my being, but also the being of the objects or persons that they are feelings for and about. They characterize me not only as a biological being, but also reveal the nature of the person that I am (whether I like it or not); or better yet, they show that my personhood is inseparably rooted in my biological nature. This brings us back to the first structural discovery about the nature of human subjectivity, namely that it is constituted by both reason and body (spontaneity and receptivity). As we saw in the first part, this was the essence of Kant's Copernican revolution. Human experience, and therefore human reality, is not a passive registration of the external world, but the product of a cooperation of reason and body. Although I agree with Strawson that the architectural elaboration of the idea is 'baroque' (Strawson 1966: 24), I think, as he himself did, that the basic idea is worth following. In fact, Strawson's seminal work on the concept of a person does come close to what Ricoeur calls the second Copernican revolution[58]. I

[58] The move from the self to an ontology of subjectivity that grounds being of the self in the being of world. This is carried out by an ontological phenomenology which 'is in its turn a sort of conversion that "decentres" the ontological concern from the Self" (PW 233/86; cf. FN 32/35, 446/448).

think that Ricoeur dismissed Strawson's analyses too quickly[59], and that a further clarification of personhood can be found exactly in his work. Strawson was, as known, deeply inspired by Kant and, although he observed that Kant's separation of the cognitive and practical side of our nature was too strict and categorical (idem: 272), he nevertheless maintained that human experience depends on the particular combination of human sensibility and reason. Human experience is peculiar and must not be identified with the notion of a detached and impersonal reality: 'The bounds of the real. We may say, are indeed not co-extensive with the types of sensible experience we in fact enjoy. We must not suppose that the nature of reality is exhausted by the kinds of knowledge which we have of it' (idem: 267). He is interested in the nature of human peculiarity, and therefore his analyses of human experience focus on the basic structures and concepts of such an experience. On line with the Kantian esthetics, he individuates time and space as the general structure (Strawson 1959: 22-3) in which material bodies 'secure us a one single common and continuously extendable framework of reference' (idem: 54). What is of particular interest to us here is that in this common framework of reference, the concept of person has a particular status because it 'is logically prior to that of any individual consciousness. The concept of a person is not to be analyzed as that of an animated body or of an embodied anima' (idem: 103). Human experience has two basic and common points of reference, namely material bodies (things) and persons. They are both fundamental in the constitution of our understanding of the world. Personhood is not the same as selfhood, and 'it is easier to understand how we can see each other, and ourselves, as persons, if we think first of the fact that we act, and act on each other, and act in accordance with a common human nature' (idem: 112).

Now, Strawson's individuation of the concept of a person as a constitutive element of human experience is important to the conception of care as the ontological ground of subjectivity because it tells us something about the complexity and, at the same time, structural unity involved in the notion of care. Our experience of, relation to, and interaction with the world, the other persons, and ourselves are rooted in care for the world, the other, and ourselves. This care involves the heterogeneity of our values that are themselves both bodily and spiritual. I care for my health,

[59] Just as Ricoeur, Strawson is not mainly interested in the objective validity of experience, but in 'ordinary human experience' (Strawson 1959: 88), and although 'persons in Strawson's sense are more on the side of substance' (OSA 83/104), it does not mean that 'that the agent as such is not yet a self, in the strong sense I [Ricoeur] am attributing to this term, but only as one of the "things" about which we speak (idem). The emphasis in Strawson's analyses is on the basic nature of personhood in our most dispassionate experience of another human being. This person is a self, with all the fragility that Ricoeur attributes to that term, but Strawson is more interested in showing that the personhood of that self is as objectively firm and inalienable as the material nature of the self. For Strawson, as for Ricoeur, the fragility of the human self is accessible on a practical level, that is, in considering the self as an agent, because 'our practices do not merely exploit our natures, they express them' (Strawson 1962: 80).

career, ambitions, economical security, my love ones, ketchup on my fries, Brague's strange trees, the hope for freedom, equality, a better world, and so on. Therefore, care is a very broad and confuse concept that may at first seem completely useless because it encompasses every thing and nothing. Nonetheless, there is a structural unity in care that makes it a bit easier to handle and therefore use in analyses of human subjectivity. Human care is bound to the structural nature of human experience, of which we have seen that the concept of a person is a constitutive feature. The subject's care for personhood is constitutive of its being, because that is what defines the 'common roots in our human nature and our membership of human communities' (Strawson 1962: 72). The effort to maintain our personhood (and not just be a selfish bastard or a moral idiot) is inscribed in most of our actions by the simple fact that our interaction with other people is constituted by practices that build on a common conception of personhood. My recognition of the other and the other's recognition of me as a person is a presupposition for our existence in a world that is constituted by me *and* the other persons. However, personhood is not only constituted by such a mutual recognition. It also involves the sense of selfhood. The person is an individual other than a social being (Strawson 1959: 112-5). My concerns are marked by this twofold nature of care. I care about the other and about myself (often more about myself than the other). This was what Ricoeur stressed when he individuated two principle desires (vital and spiritual) in subjectivity and argued that the (affective) fragility of human nature originates in the conflictual constitution of its being (body and reason). He concluded that this non-coincidence, this conflict, could only be dealt with in the pathetic dimension of subjectivity, i.e., the concrete acting and suffering of human nature. Our affective fragility, that is, the heterogeneous nature of our care, becomes a practical problem in how we deal with normative conflicts, problems, and dilemmas. An ontology of care explains the being of the human subject better than a strict materialist-inspired ontology that analyzes human subjectivity in terms of either disembodied reason (functional computation) or impersonal embodiment (evolutionary theories), since accepts the fact that the subject is rooted in biology that it can never completely control or understand (basic affective values), and still it makes room for the other fact, namely that the subject also care about being a person and being part of a human community (humanity). It is the result of what Strawson has named a 'descriptive metaphysics' (1959: 9) in that it is grounded in conceptual and experiential analyses of what it is to be a human subject.

Conclusion: Subjectivity and Ethics

The subject is a biological being *and* a person. In fact, we cannot separate one aspect from the other. A phenomenological analysis of subjectivity can only distinguish between the two by disregarding the emotional aspect of experience. This, as we have seen, cannot be done unless we strip the subject of what makes it a human subject, namely the feelings that qualify experience as my experience of things that matter to me. If we want to understand what is particularly human in subjective experience, we must include feelings in our analyses. And once we do this, personhood becomes a constitutive element of subjectivity. Things matter to the subject in virtue of its personhood. Despite this factual quality of personhood, being a person is more a problem than a fact. Subjectivity is normative in nature, and personhood makes this normativity a factual problem in the existence of the subject. Normativity is basic in the sense that it is grounded in the primitive feeling of embodied pleasure and pain. We seek pleasure and avoid pain before we think of what pleasure and pain is really about. There is always a strong element of involuntary passivity in our normative experience of good and bad. We do not choose normativity. It comes with living. We could call this the first passivity involved in the embodied nature of subjectivity. Personhood, however, accentuates the normativity of existence, since our actions and sufferings are bound to personhood, our own and that of the other. My life is not just about the thriving of my selfhood. My feelings reveal the basic nature of my being in that they orient my experience according to the heterogeneity of values. This is the second passivity of subjectivity involved in my being a person. My care in not confined to what I care about. My care is constituted by my coexistence with other persons; it involves their being, their happiness, and their suffering, as a qualification of my existence, since I experience the world as a practical space of reason configured by the values that do not have their origin in me. Alterity is expressed in both my selfhood (body) and in my being a person (the other). Subjectivity is not constituted only by selfhood, which becomes evident in the feelings of the subject. The subject also cares about being a person; or said differently, the subject feels ethical. Feeling ethical, my feeling of being a person among persons, is a basic constitutional feature in subjectivity, since this feeling, or rather diffuse amalgam of feelings, is affects profoundly the actions and the state of being of the subject. The conflictual nature of subjectivity originates in this complex nature of its being: subjectivity is a relation between selfhood and alterity by the fact of being a person embedded in an impersonal nature.

The answer to the question of the relation between subjectivity and ethics must consider the two sides of that question. What has ethics to do with a theory of subjectivity? And, does the notion of

subjectivity have implications for ethical theory? I have partly answered these questions with my analyses in this chapter, but for the sake of clarity I will conclude by articulating my considerations in two final formulations. First, the relevance of ethics for a theory of subjectivity.

The ethical nature of subjectivity is expressed in the pervasiveness of normativity. A theory of subjectivity cannot neglect this. The experience and being of the subject is inherently intertwined with the experience and being of other subjects. We cannot complete our analyses of subjectivity without clarifying that complexity. And, what is perhaps more pertinent, our structural analyses must consider that this normative feature affects and qualifies every aspect and capacity of the subject. Sensation, perception, feeling, knowledge, will, and action must all be analyzed on the background of normativity. We cannot understand subjectivity as ontologically neutral. This is not to say that structural and functional analyses that do not consider the ontological nature of the subject are unwarranted; on the contrary, they may help us understand better the ontology that grounds those structures and functions, which is in a great merit in itself. However, they cannot be said to explain the nature of subjectivity, the sense of being a human subject. Further, the phenomenological distinction between minimal self (self-awareness, ownership, and agency) and personhood (Zahavi 2005: 129-32) becomes somewhat complicated and blurred when we consider that the experiential dimension of subjectivity as a matter of fact involve the complex notion of feelings. Although we can distinguish between personhood and an experiential core self by talking about a narrative (cognitive) sense of self and an experiential self (Gallagher and Zahavi 2008: 205), this is not the only possible approach. Feelings are part of the experiential core self, and the normativity experienced by means of the feelings is personal in the sense that it reveals personhood in the experiential self. The self is feeling ethical even in the most stripped sense of subjectivity.

This brings me to the other side of the relation. Subjectivity is relevant for ethical theory in many ways. Human beings are subjective in nature. We might even say that ethical theories are necessary simply because humans are beings with individual subjective natures. But here I will focus only on the affective dimension of subjectivity because that often seems to be what divides the waters in ethical theory. The foregoing analyses have emphasized the irreducible affective nature of human values and practices. Emotions are fundamental in the ethics, and in order to understand why the subject thinks and acts ethically, we must 'focus on the emotions – on, that is, the emotions as affective, as emotional' (Stocker 1996: 124). This is basically the conclusion of my analyses. A thorough analysis of the structures and capacities of subjectivity reveals the pervasive nature of affectivity. The subject experiences and acts in a practical space of reason that is revealed by the

emotional import of the feelings. Things and persons matter to us through our feelings, and to disregard these feeling for 'cold' reasoning is to misunderstand the complexity of human affectivity. Our feelings reveal our values; and our values are expressed in what we care about, our concerns. Any ethical theory has to accept that this is a basic fact of human nature, which can be ratified by both dispassionate (neutral) observation of human behavior and structural analyses of subjectivity. Human feelings do not only involve my (egoistic) self, since the nature of human affectivity includes personhood as a basic element in emotional experience. We care about being persons, because our interaction with other persons is permeated by ethical feelings other than selfish and perverse one. This is a fragile combination of feelings, which is rooted in the vulnerable constitution of the subject. But we do not secure ethical behavior by excluding the fragility of the feelings, because feelings are not something that we choose. In our feelings we find the vulnerable constitution of selfhood (voluntary) and alterity (involuntary) in our being persons. I proposed a deflated ethical theory that considers ethics to be a matter of practical reasoning and not about *morality*. We have to find the ethical concepts that guide our behavior in what we care about in our practical life, and not in a 'moral space'. I argue for a naturalistic conception of both subjectivity and ethics that find the values of human life in the pervasive normativity that characterize human nature. Care, feelings, practices, and personhood are the key concepts in such an interrelation of subjectivity and ethics, and the foregoing analyses have only been a clarification of these concepts, not a systematic employment.

The final part of the dissertation is a further clarification of what I mean by a 'naturalistic' conception of subjectivity and ethics. I do so by examining the contemporary form of naturalization that goes under the name of neuroethics. This is a very relevant example of what I believe that naturalization should not be. My criticism is carried out on the background of the previous analyses in the sense that I look at how this form of naturalization uses the concepts of normativity, values, subjectivity, and personhood. I maintain that the central problem of this kind of naturalization is its complete lack of a coherent theory of subjectivity. This will hopefully clarify my view on the relation between subjectivity and ethics. Secondly, I clarify my conception of a naturalization of human personhood by looking at different conceptions of what we mean by nature and distinguishing my position form a recent attempt to use the process of naturalization to argue for moral relativism.

Part Three

The Brain and Human Values

Ethical behavior is a fact just as delicious red apples or an annoying headache. We all have a pretty clear idea of what is ethical and what is not. We do not think of ethics when we see a woman stumble and fall in the street. Bad luck, distraction, lousy pavements, clumsiness, and so on, but not as something that is wrong or mean. However, our reaction (or lack of same) to the situation falls under what we consider as ethical. I might run to help or think that it is not my problem that she is clumsy or unlucky. I have a train to catch, work to do, money to earn, buy a pair of shoes or met someone that I care about. Ethical behavior shapes the world as we know it, and ethical values influence how we experience situations that we find ourselves in.

I have argued for a deflated ethical theory that understands ethical values and behavior as a natural part of the constitution of the human subject. Ethical values are part of nature as we humans experience it. Ethics is not something that we superimpose on human nature or on an otherwise non-ethical world. Ethical values simply come with experience of and interaction with the world and other subjects. I have further argued that this might be accounted for by considering subjective experience configured as a practical space of reason. Our experience is structured by how we value and interact with the world by means of human practices. On this view, ethical theory is an attempt to organize, and perhaps improve, the values that we naturally experience in virtue of being human persons. So, if ethical values are a natural and factual part of being a human person, then they must reveal something about human nature. I think that our subjective feelings of values show that the human subject is a normative being that cannot be clarified without considering this normativity at the heart of subjectivity. At the same time, I claim that such a conception of human subjectivity is not to be considered a hindrance to the attempt to naturalize human personhood; on the contrary, a clarified notion of the normative structure of subjectivity can be a help for the process of naturalization. However, the process of naturalization has to accept the peculiar nature of human experience and action, and not try to make it into something that it is not. Therefore, in this third part I will try to clarify my naturalistic position by criticizing two different attempts to naturalize human personhood, namely neuroethics and a recent argument for moral relativism on the background of strong empirical naturalism. Further, I will clarify what I mean by the concept of nature and naturalism. This part is shorter than the two previous ones. This is due to the explanative nature of the presentation. It is meant as a clarification and not as a further argument, and will therefore build on the arguments presented in the two more substantial parts.

Chapter One
Neuroethics and Subjectivity

A clarification of the relation between subjectivity and ethics is particularly important today, since subjectivity can easily be forgotten in the sensationalist clamor surrounding some neuroscientific results and hypotheses concerning human nature. In recent years there has been a growing tendency, in neuroscience as well as among philosophers, psychologists, and sociologists working within the framework of 'neurophilosophy', to deal with the question of ethics. What was first sproradic attempts has now grown into the founding of a new multidisciplinary research field named 'neuroethics' (ratified in 2002 with a big two-day conference in San Francisco; see Markus 2002). Neuroethics is a distinct branch of bioethics whose 'mission is to examine the ethical, legal, and social implications of neuroscience today' (Illes 2006: IX)[60]. This is a large ground to cover, and the discipline is only just getting started, but the expectations are bold and high: 'We can improve our understanding of the nature of moral theory and its place in moral judgment if we better understand just what morality consists in. Such an understanding will best be informed by treating morality as a natural phenomenon subject to constraints from, influenced by, and ultimately reduced to the sciences, particularly the cognitive sciences and biology. Treating morality as a matter of proper function, biologically construed (e.g. at least partially fixed by our evolutionary history), with a concomitant emphasis on skillful action in the world, will also shed light on just what kind of creatures we must be (cognitively speaking) if we are to possess knowledge about morality so taken' (Casebeer 2003a: 3; see also Churchland 2006: 3; Pinker 2003: 270; Metzinger 2006: 69; Farah 2005: 38-39; Greene 2003: 849-50).

The attempt to reduce morality to natural facts is not a new one. It has been a central issue in moral philosophy at least since Hume dismissed any attempt to drag an 'ought' out of an 'is'. Moreover, ethical naturalism was considered a closed chapter after Moore, in the beginning of the last century, carefully pointed out that the question of what is good (or bad, for that sake) remains an 'open question', that is, it cannot be settled by scrutinizing natural facts. Nevertheless, with the newly gained strength and support of neuroscientific results. Ethical concepts such as morality, value, and norm now seem to be rooted in the natural soil of good, solid science. The new techniques and the firm progress of empirical science have finally made it possible to substitute the 'first guesses,'

[60] Neuroethics is divided in two main fields: the ethics of neuroscience and the neuroscience of ethics (Roskies 2002: 21). Here I will deal only with the latter.

'stories,' and 'tales from the past' told by Aristotle, Descartes, Hume or Hobbes (to put it with the words of the prominent neuroscientist Michael Gazzaniga (2006a: 164, 178)) with serious scientific insight capable of dismantling our erroneous image of human nature. Of course, this new image is met by heavy criticism from the humanities in general and philosophers in particular; therefore, when neuroscientists debate ethical issues with philosophers and ordinary laypeople they must be aware, according to the just quoted Gazzaniga, that 'No one has told the kids yet there is no Santa Claus' (idem: 163).

This new biologically informed ethics is a direct consequence of the neuroscientific image of man. Ethical naturalism has once again become a serious issue, because our knowledge of human biology has exploded with new insights into the hitherto firmly concealed 'black box' that we know as the brain. Around fifty years ago, behaviorism was criticized, because it sought to bypass the brain by relying only on observable behavior and thus deliberately ignoring the cognitive capacity of human beings. Today, on the contrary, the 'black box' has become transparent by means of various imaging techniques, and neuroscience can, allegedly, pin down the neural correlates of almost everything, cognitive as well as affective capacities, from self-referential reflection over deliberation to emotional experience (Roskies 2006: 18). Hence, it follows as a natural consequence that neuroscience also has something to say about ethics, since it gives us the possibility to register the subpersonal components of our ethical life and practices. Finally we could say something concrete and exact about this annoyingly impalpable and conflictual dimension of human nature. Some neuroethicists seem think that all we have to do is to show how the non-normative, empirically verified, structures and dynamics of neuronal networks come together to produce the normative experience that we, as human beings, have of ourselves, the other, and the world. Thus, 'the beginnings of a new paradigm in ethics can be seen emerging. Owing to the natural and biological roots of morality, this new approach to ethics may be referred to as 'naturalized ethics', or more simply, 'as neuroethics'' (Churchland 2006: 3).

Now, why include a newly formed discipline such as neuroethics in a discussion of the relation between ethics and subjectivity? I do this for three reasons.

The first of all, I do not think that neuroethics delivers what it promises; that is, it does not explain neither *why* people act the way they do and nor *how* they evaluate when they act ethically. Furthermore, it does not explain *what* morality is or how we come to think of ethics in the first place. One of the main reasons for this, I think, is that neuroethics completely bypasses the notion of subjectivity in dealing with ethics. Just like behaviorism excluded everything mental, neuroethics

deliberately ignores the first-person dimension of ethical experience. It relies exclusively on a combination of subpersonal data and environmental factors, and upon these it advances hypotheses that do not consider the dispositions, feelings, and deliberations experienced by the subject who is registered to undergo these various stimuli. Subjective experience of, for example, being a particular person with certain desires and wishes is generally considered as secondary (Zoloth 2006: 68) or even as an illusion (Farah 2007: 40). This is a problem: Can we account for human ethical dispositions and morality without paying attention to human subjectivity? Is the new, naturalized paradigm that neuroethics is supposed to provide in ethical debate actually a legitimate one?

Secondly, in my critique of neuroethics, I hope to strengthen my argument for the necessity of theory of subjectivity in ethical theory. Neuroethics emphasizes that human nature is critical to ethics. I agree perfectly with this claim. In order to make a legitimate ethical theory, we must understand how people come about to think and act ethically. Ethics requires an attentive metaethical investigation into the normative dimension of human nature. Such an investigation depends on neutral and patient observation of how values and norms are generated in human action and experience, and how they influence perception of and action in the world. I think that it is important to distinguish between neuroethics and neuroscientific research. I believe neuroscience can tell us much about human nature that is important to ethical reasoning. What I oppose is the tendency in neuroethics to go for reductive conclusions about how humans *should* behave and think according to neuroscientific findings. There often is a silent transition from descriptive analyses to normative conclusions, which impair the legitimacy of their so-called 'ethical facts'.

Finally, I surmise that by relating the previous analysis of subjectivity and ethics to neuroethics I might clarify my own naturalistic position. How can I be a naturalist and still not accept the arguments of neuroethics? As I said in the introduction, ethical dispositions such as moral sentiments (e.g. love, care, anger, benevolence, and desire) and principles fit badly into the scientific image of the world. They are considered strange and suspicious phenomena in a world made up by disinterested, swirling masses of particles. Nonetheless, we cannot deny ethics a place in this picture because our being depends on these strange phenomena. Blackburn puts it like this: 'we nearly all want to be naturalists and we all want a theory of ethics' (Blackburn 1998: 49). Now, in the attempt to place ethics in the non-ethical nature of the world, we might not just learn something about the nature of ethics, but perhaps ethics might also clarify problematic aspects of some versions of naturalization. The practical necessity of finding an appropriate ethical theory that accounts for how we actually think and act in the world makes it reasonable to presume that ethics

should have a legitimate claim on the 'big' question about human nature. Actually, the fact that we need to find room for ethics in the disenchanted world is confirmed precisely by a phenomenon as neuroethics. Even the most hardcore eliminativist wants to legitimize his or hers (im)personal conduct. Ethics stresses the problematic relation between fact and normativity. Neuroethics performs a rapid straightforward translation from fact to norm, inscribed in a larger evolutionary context. Often it goes something like this: when we analyze the constitution of the human subject we find out what it is and how it reacts in certain circumstances, therefore we can prescribe that one kind of cognitive constraint is more scientifically based than others; for example a principle of utility instead of a deontological principle (Farah and Heberlein 2007: 46). However, a good translation must pay attention to both sides of the equation, that is, both the 'is' and the 'ought'. If we want to find the best translation, we must look for a theory that includes as many dimensions as possible.

Thus, the following criticism will try to argue for the claim that neuroscientific findings must rely on more than evolutionary theories to provide an adequate ethical theory; it must pay equal attention to, for example, the dimension of subjectivity, socio-cultural factors, history, logic, institutions, and economics. Ethics cannot be dealt with only from, to say in a simplistic manner, a synaptic perspective. As Moore writes: '[if] we once recognize that we must start our Ethics without a definition, we shall be much more apt to look around us, before we adopt any ethical principle whatever; and the more we look about us, the less likely are we to adopt a false one' (Moore 1954: 20).

In this open search for an adequate ethical theory, we may have to modify the process of naturalization. An adequate ethical theory is a descriptive notion that should function as a heuristic tool in our understanding of normativity. If we accept that there exist both 'is' and 'ought' in the human world, then we should try to describe both aspects as carefully as possible in order to procure the most adequate picture of human nature. Humans are individual subjects with their peculiar ideas, sentiments, principles, and perspective on things. We cannot dismiss a person's benevolence by referring to a selfish gene that only strives for survival. And we cannot explain altruistic behavior with the fact that survival is more likely in larger groups than by oneself. Values are heterogeneous phenomena that cannot be boiled down to one single principle such as survival. Values are felt, thought, and wished for by the subject considered as an individual person (made up of idiosyncratic tendencies and socio-cultural factors as well as biologically defined traits); and this

subjective dimension of values cannot be eliminated as irrelevant, byproducts, or illusions, if we want find a reasonable account of values.

In neuroethics, the neuroscientific findings are commonly interpreted in terms of a 'neuroessentialism' that states that 'our brains define who we are' (Roskies 2002: 22), and that '[a]dvances in neuroscientific research in relevant areas may change the very fabric of our philosophical out look on life (idem). On the following pages, I will advocate that an inverse model may be needed if we want to understand what neuroscientific research might contribute to our understanding of ethics. A philosophical consideration on the nature of ethical dispositions, values and norms may be a useful for the way we approach the question of naturalization. Ethics can be seen as a counterweight to 'neuroessentialism' because it emphasizes the complexity of normativity and thus clarify the difficulties involved in translating 'is' to 'ought' or reducing 'ought' to 'is'. Further, on my own view of ethics, the problem of neuroethics lies mainly in its reductive stance toward subjectivity. 'Ought' is a subjective experience of a world that can be described as non-ethical by the methods used by the natural sciences. However, this normative feature of the world is not something that we should try to reduce to what it is not, namely non-normative facts. Metaethics is an objective approach to the subjective experience of normativity in the sense that it seeks to clarify, objectively, the nature and workings of our values, principles, and norms. Our normative experience of the world is as legitimate as our distanced and neutral approach to that world. The problem is how to reconcile the two without reducing the one to the other. This will be the topic of the final chapter.

But first I shall present some of the arguments proposed by William D. Casebeer whose book *Natural Ethical Facts* provides, to my knowledge, the most coherent theoretical background for neuroethical proposals. I concentrate on the question of values. To avoid that I give a narrow view of the discipline, I integrate some other authors in the presentation of Casebeer's arguments.

The Case of Casebeer and Some of his Allies: Impersonal Ethical Facts

Casebeer wants to explain our ethical behavior in natural terms, which in itself is not a bad thing. We must, however, understand what he means by natural. He uses three approaches to clarify his natural ethical facts: neurophysiology, cognitive sciences (especially connectionism) and evolutionary biology (2003a: 11). Although Casebeer only rarely goes into more specific neurophysiological considerations (2003b: 843-4), he nevertheless builds his theory on connectionism because this unifies neurobiology and cognitive psychology in a functional model

that, according to him, explains moral reasoning (2003a: 6). Roughly defined, connectionism is a variation of the older computational approach to human cognition. But whereas computational theories do not aspire (or even intend) to model the neurophysiological anatomy and workings of the brain, connectionism wants to do exactly that. By using the neural networks of the brain as a model, this kind of computations aspires to tell us something about the way we humans reason. There are, however, many incompatibilities between connectionism and the actual neurophysiology of the brain, since the model offers more than is anatomically possible, for example information is allowed to flow in both directions along the axon (Pinker 1999: 114). But this is not the question here. In short, connectionism is an attempt to understand the workings of the brain (mental processes such as thoughts, deliberation, decisions, ideas, associations etc.) by assuming that mental processes are computational processes defined by simple syntactic specified entities that, when they are accumulated in a recursive system, functions as representations of the world. In other words, the brain functions as computer. However, I think that Casebeer's theory (and consequently many neuroethical theories) can be presented faithfully by adopting a more general point of view, namely with regard to his account of natural values.

His basic idea goes something like this: if we connect a sufficiently large amount of neural networks (artificial or natural), we could eventually demonstrate that the more networks that are combined in an organism, the smaller the actual difference between learning and judging would be. In a recursive system, in which input and output are regulated by organism's ability to adapt to the factual surroundings, the capacity to judge simply consists in the ability to create the best possible options for existing in those surroundings: 'In a sufficiently complex neural networks – exhibiting sufficient recurrence, coping with the world, and interacting with the environment – 'comportmental' behavior would arise naturally' (Casebeer 2003a: 86). The question about *how* the organism should behave is reduced to a question about *what* is best for the life of the organism. Casebeer is very clear about the fact that his functionalist account of behavior is grounded in evolutionary biology. The life of biological organisms is characterized according to what is considered to be 'ecological valid' behaviors 'that let them adapt rules and regulations to situations in a manner that promotes flourishing' (145). He challenges the traditional view that morality is a particular human quality (75) by emphasizing that all organisms act according to values simply because 'moral facts are functional facts' (53).

He distinguishes between three kinds of moral agents (92-3): 1) minimal moral agents (plants, viruses, bacteria, and some insects), 2) standard moral agents (insects and most other animals, 3)

robust moral agents (humans and perhaps dolphins or higher primates). Humans are only different from the amoeba because of the more complex neural networks working in the human brain. This complexity allow us to have plans, projects, and desires that are not directly related to immediate or proximate satisfaction; in fact, Casebeer ventures that 'possession of numerous such projects is in fact the singular mark of humanity' (93). On the other hand, moral qualifications can be extended to viruses, because, as he points out, we do in fact say: 'Lo! A *flourishing* Virus!' (92). Thus, Casebeer prepares a ground on which we can speak of moral values just like we speak of anything else in the biological universe.

This is one of the main principles in neuroethics. If we understand the evolutionary and physiological aspect of the brain, we will be able to see 'the *very* large extent to which information processing takes place automatically, below the level of conscious experience' (Levy 2007: 20). And that which actually depends on conscious experience, our decision-making, is somehow also not to be taken personally: 'we *are* the set of our subpersonal mechanisms, and nothing else, and our intelligence just *is* the product of the successful combination of these dumb machines […] for the mechanisms that make the decision are nevertheless ours, us; they have our values, they have our beliefs, our goals (we have them *by* them having them), and when they decide, *we* decide' (idem: 243). Nature is not personal. This is obvious even to those who are not neuroscientist or biologists. Already the Great Lisbon Earthquake in 1755 diffused this austere reality among the intellectuals in that time, and in a time of wars, torture, natural catastrophes, and desperate immigration who would today maintain otherwise. Nature is heartless and indifferent. The distinction between human agency (e.g. war, terror, and immigration) and anonymous events is made difficult by such a functional account, because humans are, in virtue of being nature's progeny, deeply immerged in this impersonal matter. And to assume personhood to be part of the natural world is the same as to think that the stick is broken in the water just because we are ignorant of the physical laws of optical reflection that underlie this visual illusion. Our ideas of personhood are merely due to a 'misleading system for person representation' that should be left out of ethical debate, because 'personhood is not really in the world' (Farah and Heberlein 2007: 45).

If we return to Casebeer, we can see that just as we naturally select the most flourishing and beautiful orchid in the flower shop, a functional evolutionary ethics secures values that 'remain true to the neurobiology of moral cognition and the empirical facts about successful ways to produce human flourishing' (2003a: 137). In order to reject what he considers to be implausible norms and values, we must understand that '[n]orms to which cognition is ultimately responsive would be

pragmatic "fit-functional" norms' (90), and the reason why we humans so often end up in moral disagreements and conflicts is 'because we have constructed a faulty model that does not effectively link the demands of our functional nature to the structure of the world' (93). The fit-functional norms and values of orchids in a flower shop *are* to be flourishing and beautiful, therefore they *should* be so otherwise they do not fit the proper function of orchids and will *naturally* end in the dumpster. Casebeer does not believe that such a radical approach to human ethical behavior is in contrast with some of the great moral philosophers of the past. He claims to develop a 'robustly naturalized Aristotelian ethics' that combined with Dewey's pragmatism guards the natural nature of morality against 'allegations about the relativity and queerness of objective values' (49). Values are objective and travel cross nature's many species because if we stick to solid scientific evidence, it is difficult to distinguish what is particularly human behavior from that of a vampire bat (Kennedy 2002: 199-200).

The ontology that sustains and nourishes this picture of nature and human values is easily identified as a strong materialist ontology that emphasizes the epiphenomenal nature of our subjective experience of the world and ourselves. What counts is the impersonal constitution of such experiences, and the only hope to find real values that escape the implausible claims of both non-naturalist realists such as Plato, Kant, and Moore and anti-realists or skeptics such as Hare, Blackburn, and Mackie is to dig deeper into the constitution of human nature, and let the most sophisticated of our bodily members, our brain, speak for itself. Researcher have actually located the utilitarian center in the brain (Moll and Oliveira-Souza 2007), which seems to confirm that neurobiology may have an important role in discussion of moral judgments; for example, it can show us that 'moral motivation and moral judgment are functionally separable and depend upon distinct brain systems', which may have 'radical implications' for moral education (Roskies 2006: 29) that in the end can allow us to develop 'a more morally responsive citizenry' (idem: 30). Such insights will eventually help us 'live fully functional lives' and 'structure character-development institutions such as schools and colleges' (Casebeer 2003a: 145).

Thus, the best way to solve the conflictual nature of human society is to look for something more stabile and solid than the fragile and vulnerable nature of human subjectivity. It is, in fact, a characteristic feature of many publications in neuroethics that they very easily jump from descriptive metaethical investigations to normative moral policy, educational and institutional proposals. Neuroethics is an urgent help to the society that all too long has suffered under the hopeless and utopian regime of folklore and philosophy. Instead of spending useless time and

energy on Moore's irreducible 'ought' and on how it is connected to the factual world's 'is', we should rather understand that our biological nature, factual environment, and evolutionary theory has already given us the tools to commence 'our own new enlightenment' of ethical values (Saphire 2002: 5).

Prominent philosophers such as Dennett, Metzinger, and Paul and Patrizia Churchland attempt to provide the young discipline with philosophical arguments against moral philosophers who claim that such ethical reasoning is unfounded or simple wrong. Dennett, for example, recommends that we substitute our elusive ideal of genuine altruistic behavior with a more natural '*benselfishness*' (Dennett 2003: 194)[61], which is found in all species and confirmed by evolutionary research. Therefore instead of chasing Kant's deontological fantasy (idem: 213), we should put a lot of computers to do 'simulations of evolutionary scenarios' of possible actions and find out what would be the best behavior in the future (idem: 218).

The Disobeying Subject: Subjective Values and Personhood

'His thoughts wandered again. Almost unconsciously he traced with his finger in the dust on the table: 2 + 2 = 5 […] But it was all right, everything was right, the struggle was finished. He had won the victory over himself. He loved Big Brother' (Orwell 1989: 303, 311). These are words from the concluding pages of George Orwell's novel *Nineteen Eighty-Four*. After having secretly fought Big Brother's inhumane norms and values (peace is war, truth is manipulation, love is torture, sex is impersonal reproduction) with his small personal revolts such as constantly thinking about the history that it is his job to rewrite, drinking too much gin, turning his back to the telescreen and making love to Julia, Winston Smith finally, with heavy help from brutal prison guards in the Ministry of Love, reaches to the conclusion that they were right and he was wrong: 2 + 2 is indeed 5.

Now, I do not want to insinuate that neuroethics aspires to the terrifying account of impersonal functionalism that we find in Orwell's book. I do think, however, that by rejecting the subjective aspect of human nature we end up with a distorted account of norms and values that disregard the complexity of human nature. The values and norms of humanity are reduced to the impersonal forces of genes, synapses, distant evolutionary drives, and fit-functional societies. To argue that we

[61] Dennett has, with his usual astuteness, coined this neologism from an old anecdote, which tells that Benjamin Franklin said the following words to John Hancock, at the signing of the Declaration of Independence: '*We must indeed all hang together, or, most assuredly, we shall all hang separately*' (Dennett 2003: 1993). He plays with the double meaning of 'Ben –': Ben(jamin) and Ben(e) as in the Latin word for good.

cannot do ethics like this, I will look more closely on the notion of value as it is conceived and employed by at least some of the spokesmen of neuroethics, and argue that we cannot hope to solve the problems involved in this notion without paying attention to the subjective nature of personhood.

Casebeer notices that he does not distinguish between ethics and morality, because, although he obscurely agrees that such a distinction 'might be eminently useful' in other contexts, it is not relevant for his project to naturalize ethical/moral reasoning (2003a: 13). Although I agree that a sharp distinction is futile, I do believe that sometimes such a distinction can prevent confusion with regard to another distinction that we cannot permit ourselves to confuse, namely that between factual description of normativity (metaethics) and normative claims (moral debate). Metaethics deals with the nature of moral value, moral motivation and the status of moral judgments, whereas moral debate can be seen as the (important) discussion of actual rules and norms (written and unwritten) of a given person, society, or even humanity in general. The one is not more important than the other, but they involve different kinds of analyses. Metaethics begins without any normative claims and attempts to clarify the nature of the claims that are being discussed in moral debate. I think that neuroethical analyses often confuses the two because they built their analyses on a normative claim that facts and nature are reducible to what are revealed by the empirical sciences and evolutionary theory. There is no doubt that they are right a long part of the way. Neuroscientific research can help us understand many relevant facts about how our brain works, and this is necessary information that any metaethical analysis has to pay due attention to. But it is simply wrong to use neuroscientific results to justify blunt statements like: 'moral facts are functional facts' (Casebeer 2003a: 53); or that we must opt for a utilitarian ethics because personhood is an illusion (Farah and Heberlein 2007: 45); or that morality depends on intuitions because 'moral judgments are not rational' (Levy 2007: 293) and therefore 'perhaps some kind of consequentialism is the best normative theory' (idem: 306); or that '[t]he general cognitivist picture of morality is correct: moral judgments are a subspecies of judgment; moral reasoning is a type of reasoning; moral concepts are an ordinary type of concept, and there are nothing intrinsically motivating about either moral concepts or moral judgments' (Roskies 2006: 28). The reason why it is wrong is that these researchers use the neuroscientific results to settle debates that involve more than neuroscience can deliver. It is like saying that just because a hammer is the best tool, if we want to fix the broken wood frame of our favorite picture, then it is also the best tool to fix our ragged marriage. Neuroscience is a good tool, but it is not the only tool to cope with human nature and values. The

problem is that neuroethics is, per definition, grounded in an ontology that have a strict normative criteria for what is real and what is not. I return to the discussion of the concept of nature in the next, concluding chapter, but for now I limit myself to the nature of values.

For neuroethics, real value does not depend on subjective experience. Subjectivity is hardly ever mentioned or discussed in the literature, and when it is it is not considered as a first-person perspective on the world and the other in their phenomenological givenness, but in relation to constructivist accounts such as narrative ethics (e.g. Zoloth 2006: 68-9; Levy 2007: 175-7). If we dismiss the subjective nature of the agent, how do we then decide on the nature of values? The answer is straightforward: the good values are the real and true values that must be discovered and ratified by objective empirical sciences. This will then show us what is best for human beings according to their biological nature and the function of social interaction. The subject does not have anything to say about the values by which it leads its life, because its whole existence is spun up by the illusion that personhood counts in human interaction. Whereas persons experience doubts and confusion with regard to values (what are my values? How do I know if I adopt the right ones? Why are my values better than the idiot who sleeps until noon every day of his useless life?), the neuroscientifically good values are stable and clear simply because they 'are discovered by examination of the biological world of organism-environment interaction: they are facts, empirical matters in any reasonable sense of the phrase' (Casebeer 2003a: 33). Casebeer, who is one of the few neuroethicists who actually deal with Moore's open-question, rejects the possibility that the nature of values is more a problem than a fact, since we all have different attitudes, opinions, and feelings concerning the values that govern human interaction. He is convinced that if we let the nature of value be an open question, then 'it can legitimize armchair metaphysics' (idem: 32), which will eventually hinder any 'moral progress' and 'fruitful lives' (idem: 161).

Even without considering the logical aspect of Moore's question (the analytical nature of the notion 'good'), we easily see how such a functional account of values goes against everything that we normally understand by human values. Values are a factual problem because individual persons actually do have different values that sometimes contrast, sometimes concur, and sometimes start wars. A part of the job of metaethics is to investigate into the nature of these different values and try to argue rationally for why some values are more in accordance with what we know about human nature, the nature of the world, and human interaction. In (most of) the western world today, homophobia, repression of women, child labor, extreme nationalism, death penalty, race discrimination, and violence are criticized because such behaviors and attitudes express values that

do not respect the sovereign right of the individual person. The person is an individual person and has a right to his or her individuality (feelings, opinions, principles, dreams, and attitudes). Ethics comes into the picture because often this individuality compromises the individuality of other persons. And metaethical theory is the attempt to clarify the concepts and arguments we use when we think, act, and discuss ethically.

Our biological nature, as it has been uncovered by the empirical sciences in the last 150 years, is a functional organism that is born and developed by specific (although many yet unknown) laws and processes. These sciences have enabled us to dissipate many absurd myths and urban legends about human nature (there is no such thing as racial inferiority, pale skin is not more noble than dark skin, women are not made in the image of man, we are not the center of the universe, left handed people are not just being obnoxious, greed is not only Jewish, and tuberculosis is not cured by fresh mountain air). Functional analyses of nature (human as not human) do a lot of good. But although the normative dimension of existence can find precious help in such analyses, it remains a categorical mistake to think that values can be explained or even characterized as mere functional facts. This is to explain and instruct human nature by making it into something that it is not, namely impersonal. We saw in the discussion about the evolutionary approach to emotions that intentionality cannot be explained by distant evolutionary forces, since human beings care about individuals (H-intentionality) in difference to other mammals. Our rationality is governed by something more than efficiency and functionality. Most of us do find it wrong to sacrifice an individual to save thousand human beings. And it is a misunderstanding of rationality to say that it would be more rational to do so, because then we save many other lives. But in this way of thinking, it would perhaps be more rational to save one, since the earth is becoming overpopulated, which is a bad thing, rationally speaking, because then humanity will eventually die of lack of oxygen, food, and water. Our rationality does not work in a vacuum, but is strictly interconnected with the rest of our human nature, that is, our feelings, sensations, memory, and cognitively impenetrable desires. And one pertinent feature of human beings is that they do care about being individual persons with all that this involves. It was De Sousa who in last chapter brought attention to this feature of human evolution by emphasizing the particular individual nature of human intentionality. He has recently written an interesting little book about the evolutionary development of rationality and individuality where he points to the fact that '[w]ithout the power of thought, whatever I do and whatever happens to me, the ultimate *why* of my behavior can only reside *outside* me, in the vestigial teleology of my genes […] By means of the explicit resources of language,

humans became conscious of themselves as individuals and empowered to invent new values transcending the vestigial values embedded in our biology' (de Sousa 2007a: 86)[62]. Ricoeur's conception of values stressed this complexity in the nature of what we value, of what we care about, by defining value as a quasi-concept. The concept of value encompasses many heterogeneous aspects about human care or concerns that we cannot understand, only feel. Values are personal, impersonal, spiritual, biological, historical, traditional, uniting, separating, universal, parochial, firm, changing, fragile, pure, nasty, transparent, obscure, and so on. Human nature is expressed in the values by which we interact with one another. We can never completely decide on what to care about, although we can reflectively learn to control some of what is not appropriate or right to value. The value of erotic love is a good example. Blackburn writes: 'a partner who can decide at will whether to feel desire is not quite the real thing. We don't want control. We want to feel swept away ourselves, and especially we want each other to be swept away, just as we require blushes to be involuntary, and it is no sign of shame that they are' (Blackburn 2004: 61). The experience of being bound to our values is an expression of the cognitively impenetrable nature of what we care about. There are involuntary aspects of values that cannot be disregarded without ending up talking about what is the right and wrong values (i.e., normative claims), and not about the nature of values. I know that Peter is bad company but I cannot help wanting to be near him. It is like explaining homosexuality with reference to the fact that nature (and your mother) wanted it otherwise; or, that we should not use preservatives because sex is just reproduction and to hinder reproduction while

[62] Another recent and important publication is the collection of essays by neuroscientists, psychologists, and philosophers: *Tall Tales about the Mind and Brain. Separating fact from fiction*. The book is meant as an attempt to dissipate the myths and legends concerning neuroscience, both among the researchers inside the discipline itself and on part of the public. For example, with respect to the present discussion, the neurophysiologist, Giovanni Berlucchi, argues that it is hard to dispose of individuality for impersonal function even at microscopic level, because '[a] person's identity, individuality is the result of that person's genome as much as of his our hers unique and unrepeatable life history, including prenatal development in the intrauterine environment. Further, in addition to genetic and environmental influences, there is an independent third source of phenotypic variation which has a major role in ensuring the uniqueness of each individual organism and particularly of each individual brain. Independent of their genetic make-up, persons have different minds arguably because the have different nervous systems, and they have different nervous systems because each nervous system is the unique, irreproducible result of three factors: the genes, the environment and developmental noise' (Berlucchi 2007: 337). This can be used against the arguments that some neuroethicists use when they want to defend themselves against accusations of biological determinism. They often state that a person's behaviour is not only in his or hers genes, because '[c]ontext is everything – and this, quite simple, is the lesson of neuroscience' (Gazzaniga 2006b: 143). A person is his or hers biological constitution *and* the environment. But this is to miss the point, namely that a substantial amount of neuroscientific studies show that our individuality depends on, at least, a third feature, developmental noise, which is best explained by the fact 'that chance events occurring during ontogenesis contribute to the production of unique body/brain composites, each of which will interact with the environment in its own idiosyncratic manner' (Berlucchi 2007: 345). An individual person is more complex, even at a microscopic biological level, than just the combination of genes and environment. Therefore we cannot rely on biology and a good functional society to envisage the functional character of human nature. Every human being functions in his or hers unique way.

having sex is irrational. To understand and explain what human values are is actually to explain much of what is to be a human being, and vice versa. When some researches claim that the problem with human beings is not that they have too little morality but too much (Pinker 2003: 269), then this reflects a rather undisturbed idea of what human nature really is. Values are approached with a very clear idea of what is human and what is not. We should not think that morality is a complex and multi-layered question, because '[t]he modern sciences of mind, genes, and evolution are increasingly showing that it is not true' (idem: 421). The only problem is that people do have so much morality and care about what they are not supposed to care about, and unless we use the same 'well-meant' methods as the Ministry of Love, we will never change people's attitudes, feelings, and rationality in such a dramatic way that 2 + 2 will eventually become 5. Our values spring from our subjective experience of the world, and, as I have been arguing for all this time, this experience is a fragile constitution of the voluntary and involuntary aspects of human nature, selfhood and alterity, spontaneity and passivity. And when, for example, Casebeer writes that '[w]hether or not certain persons are irredeemably evil is an empirical question, amenable to inquiry' (2004: 450), I would not disagree that this is in part true. Evil is empirical. It is a tragic fact about human intercourse and society. Too many people suffer from mischiefs so cruel and horrifying that we cannot but consider the persons who do such things irredeemably evil. But this is only part of the story. The real difficulty arises with our experience of evil. How and why do we experience evil? What makes us characterize actions and persons as evil? It is the experiential dimension behind the empirical fact, the values that structure our normative responses, which is more complex than expressed in the biological-functional account of values and facts[63].

[63] Ricoeur's work on the symbolic nature of evil remains an important source of clarification with regard to the complex relation of language, concrete sedimentation of daily routines, myth, and religion in our experience of evil. When we look at the experiential dimension of evil '[t]hen the problem will arise, how the quasi-being and the quasi-nothingness of human evil are articulated upon the being of man and upon the nothingness of his finitude' (SE 357/488). Evil is a fact. However, this factuality continuously subtracts itself from our comprehension because of the nature of our experience of evil, which comes in endless forms, degrees, and physical manifestations. Evil is a problem in (human) nature. And the real problem is how to confront that evil and how to levitate the suffering that evil brings about. In order to do so, we need to understand the fragility of our knowledge of evil. Is it outside of me, part of me, part of the world, in the other, in the traces of tradition and religion, and can we wash it away like a stain on our conscience? Are my values, my care, causing evil without my intention to do so? If we look at the experiential dimension of evil, then this 'irredeemable' fact loses its empirical clarity and works its way into our own personhood. If I realize that I cannot separate myself from evil, maybe I will come to 'understand that evil is not symmetrical with the good, wickedness is not something that replaces the goodness of man; it is the staining, the darkening, the disfiguring of an innocence, a light, and a beauty that remain. However *radical* evil may be, it cannot be as *primordial* as goodness' (SE 156/306). I never consider myself 'irredeemable evil', and if I see the evil in my own being, perhaps I will see that it is not so much a fact about evil persons as a problem about our being human that 'points toward the relation of radical evil to the very being of man' (SE 156/305).

The relation of ethics and values demands more than subjectivity. This is obvious. Ethical values cannot just express the subjective nature of personhood, since we all have different values that have to be somehow reconciled if we want to live together in a world that we all share. Therefore, functionality should not be considered a bad word in our ethical reasoning. On the contrary, human values must reflect the demand of a certain degree of functionality which I have tried to show earlier by emphasizing the practical nature of our experience of and relation to the world and other subjects. However, I do not believe that we reach this degree of functionality by excluding the subjective nature of our being. Without a respect for the subjective character of human personhood, the notion of functionality can end up destroying the person that ethical reasoning wants to protect. Fact and value are intrinsically interwoven because the human person is constituted by both, and as such they cannot be divided in what is fact and what is value, or value reduced to what is fact, or fact excluded from value. To say that to assess and understand evil or good is basically an empirical inquiry is to decide on what is good and evil before starting the inquiry, which normally would be considered a wrong way to proceed.

A similar kind of reasoning sometimes affects the legal system. The political philosopher Brian Barry criticizes the functional logic of incapacitation of habitual offenders that haunts the U. S. legal system (Barry 2005: 95-105). The so-called 'Tree-Strikes-Out' laws permit up to life sentence for a third felony conviction. These laws are supposed to incapacitate habitual violent criminals and thereby make the society more safe, but have in longer periods resulted in life-sentences for twice as many marihuana user than kidnappers, rapists, and murderers all together. And in 2000, a black man in Texas was sentenced to 16 years imprisonment for theft of a 1 $ Snickers bar. Explaining the hard sentence to the public afterwards, the district attorney said that besides the fact that he was a habitual offender this wasn't just a candy bar, but a king size candy bar (idem: 99). This was of course meant as a very bad joke, but the fact remains that 85 % of the convictions made by the 'three strikes' laws in the 1990's involved non-violent crimes, often minor misdemeanor.

The functional nature of legal systems is obviously different from those governing ethical conduct, but if we look at the frightening procedures of many legal cases, we might get an idea of how strict functionalism lead to inhumane treatment of human beings. When we forget (or purposely disregard) that the accused is an individual person and instead see him or her as a certain type (habitual offender, poor, male, black), then our action or final sentence will necessarily lose what normally qualifies it as human (that we care for individuals).

We should not, however, delude ourselves by thinking that this functional logic is restricted to an unjust (North American) legal system. Another political philosopher, Giorgio Agamben, has argued that this intentional negation of the individual for the sake of the state is deeply rooted in the tradition of western politics. Ever since Aristotle characterized man as a political animal, the distinction between life common to all living beings (zoê) and particular kind of life, i.e., qualified life, (bíos) have grown more and more indistinct and the bare life 'gradually begins to coincide with the political realm, and exclusion and inclusion, outside and inside, bíos and zoê, right and fact, enter into a zone of irreducible indistinction' (Agamben 1998: 9/12). The bare life (vita nuda) is neither animal nor human life, but simply life as unqualified biological existence. Not personhood or humanity is the fundament of politics but the naked biological fact of existence. Agamben traces, in his dense and complicate writings, this tendency to define the bare life up through western ethico-political thought to what he (with Arendt and Foucault) calls the basis of biopolitics. Human nature is included into workings of biological nature by means of 'the anthropological machine of the moderns', which 'functions by excluding as not (yet) human an already human being from itself, that is, by animalizing the human, by isolating the nonhuman within the human: Homo alalus, or the ape-man […] What would thus be obtained, however, is neither an animal life nor a human life, but only a life that is separated and excluded from itself – only a bare life' (Agamben 2004: 37/43). Agamben's point is radical. He sees this reduction of human life to bare life as the criterion with which modern western democracies gain the legitimacy of their sovereign power over the individual human being in order to promote the flourishing of the state. When the right to humanity is no longer considered a fact of the individual person but a decision to be made on part of the state, then the choice is open to exclude from humanity that which is not considered human in the human being (the animal or simply the different). Once we start to politicize human life, arguing about what is human and what is not, the including exclusion of bare life becomes a reality. This is the paradox in the modern human rights: 'In modern biopolitics, sovereign is he who decides on the value or nonvalue of life as such. Life – which, with the declaration of rights, had as such been invested with the principle of sovereignty – now itself becomes the place of a sovereign decision' (Agamben 1998: 142/158). The contemporary political emphasis of the value of bare life dissolves the primordial humanity of the individual person, because the human values are questioned in the first place. It becomes possible to exclude when we want to include some traits and not others in what makes a living being human (idem: 7-8/10-1).

This may be a necessary function of a political or legal system. Ethics, however, is neither politics nor jurisprudence. Ethical values should protect the individual person from the objectifying, impersonal reasoning that characterizes many legal and political procedures where persons are to be 'managed or handled or cured or trained' (Strawson 1962: 66). In such procedures, persons can be stripped of their humanity and suffer from the 'intuition of fittingness' that, when liberated from human sentiment, is free to work out 'an adequate basis for certain social practices in calculated consequences' (idem: 79). Agamben's picture of the mechanisms of western democracy may be glooming and radical, but not distorting with respect to factual state of world politics today. Biopolitics has been the horrifying agenda for much politics in the twentieth century, and today the accelerating focus on bodily welfare, social aesthetics, and the materialist obsession expressed in consumer directed policies have a strong hold on the political agenda in the majority of contemporary western societies. This threatens to undermine the values that made moral and social progress possible in the first place. The problem with functionalist accounts of values is that they tend to forget the concrete historical consequences of past realizations of functionalism (e.g. Holocaust and the Gulags). Agamben's analyses show us that with human reason there is always the risk of categorizing and excluding that which does not 'fit' into what others tell us is right and good. Human values are fragile. And we must be careful with the theoretical (philosophical, economical, or political) conclusions that we draw from *our* account of values, because such theoretical conclusions do actually affect our own lives and that of others[64].

[64] I do not think that we should consider this a distant possibility that *might* occur or does occur in less enlightened societies than the western. It is risk that lives, and sometimes even flourishes, at the heart of our own 'civilized' values. The economist and social thinker Robert C. Reich has in his latest book argued that what he calls 'democratic capitalism' is rapidly transforming itself into 'supercapitalism', which suppresses the 'citizen' in favour of the 'consumer'. He sees this as a dangerous shift in mentality, because it threatens to undermine the humanity of social and political structures in the name of economics. He writes, for example: 'Economics, as a discipline, focuses on a domain of personal concern strictly bounded by what analysts in government statistical agencies define as one's "family" or "household." But such categories are arbitrary. The capacity of human beings to emphasize – feel responsibility, loyalty, and simple human connection – extends far beyond them […] Standard economic models have little to say about any of these altruistic sentiments. Yet as citizens we may care a great deal if most people's jobs are insecure and wages stagnant, and if a relatively small number of people have cornered most of the nation's wealth' (Reich 2007: 98). Reich mainly deals with US politics and democracy, but he stresses that this is not a local phenomena: 'People around the globe are more able to pursue their own desires and profit from their investments with increasing fervor. Yet despite the satisfaction they feel as consumers and investors, many are frustrated in their capacities as citizens. Their democracies, too, are finding it more and more difficult to articulate and act upon the common good. Voters surveys in Britain, Italy, Spain, Belgium, the Netherlands, Norway, Sweden, Ireland, and Japan show citizens who have grown to feel almost as disempowered as Americans' (idem: 9). Reich, however, thinks that it is too easy just to blame large companies and corrupt politicians for this sad change: 'You and I are complicit. As consumers and investors, we make the whole world run. Markets have become extraordinarily responsive to our wishes – more so all the time. Yet most of us are of two minds, and it is the citizen in us that has become relatively powerless. Supercapitalism is triumphant. *Democratic* capitalism is not' (idem: 99)

The neurobioethical focus on the biological and societal nature of homo sapiens tends to forget that what makes us human is not an unqualified amalgam of the pure beating of our heart, our 'specialized systems of representations, such as mathematical or chemical symbols' (Levy 2007: 43), and a fit-functional society. When Casebeer argues that he is actually just updating Aristotle with the neuroscientific insights of present day, he fails to mention that his Aristotle has completely forgotten that he once wrote the 10th book of the *Nicomachean Ethics* where he described human contemplation as the highest form of human happiness. The fundament of humanity is not the impersonal bodily material of human work, an undefined human flourishing, or the procedures of law and politics, 'for to do noble and good deeds is a thing desirable for its own sake' (Aristotle 1984: 1859 [1176b7-8]). This kind of contemplation is not functional reasoning, since it is not directed at anything else. It is the liberty of every human being to think, dream, feel, value, and understand the world in his or hers own personal way. Ethics is our way of conditioning this subjective basis of human personhood by finding the most appropriate actions according to our individual natures. If we destroy or objectify this subjective fundament of human intercourse by making our ethics functional, then 2 + 2 can eventually become 5 and the absurdity of 'Arbeit macht frei' can once again be justified by a greater meaning[65].

This critique of neuroethics should not be read as new version of the old criticism of rationality and scientific and technological progress. As I have said before, I do believe that a further scientific clarification of human nature can help us make a better ethical theory. What I oppose is blind enthusiasm and quick solutions to difficult problems. Aristotle, Hume, Kant, and Moore were all cautious, almost enigmatic, with regard to human nature. Ricoeur joined this attitude by emphasizing the fragile constitution of subjectivity. We know that human nature is a part of physical nature and still somehow autonomous of that same nature. Humans are self-centered, need company, seek happiness, and reason about how to live and how to act in the best possible ways. But to say that human beings are only reasonable or only very complex biological organisms or made in the image of transcending gods is to say more than we actually know. We can choose to say like Changeux, when Ricoeur confronts him with this fragility of our knowledge, that this will all be completely different in 2500, where we can finally rewrite Spinoza's *Ethics* on

[65] Agamben describes the contemplative nature as the fundament for selfhood and thereby human action and life: 'The life, which contemplates its (own) power to act, makes itself non-workable in all its operations – it lives the (personal) life. I write "own" and "personal" in parenthesis because only through the contemplation of power, which makes non-workable every specific *energeia*, something like the experience of an "own" and a "self" becomes possible. The self, the subjectivity, is that which opens itself as a non-workability central in every operation, as the liv-*ability* [viv-*ibilità*] of every life. In this non-workability, the life that we live is only the life through which we live, only our power to act and to live, our act-*ability* [ag-*ibilità*] and our liv-*ability*' (Agamben 2007: 274. My Translation).

neurobiological terms (WT 201/211). Then we will eventually understand that ethics is just an extended part of a larger neurobiological account of the living organism. Such outcries, however, are neglecting the fact that a neurobiological account of ethics will always remain functional 'from senseless atoms to freely chosen actions' (Dennett 2003: 306), since it is embedded in an obscure evolutionary theory that enhances 'benselfish types' and 'punish too-pure "altruists"' because these 'help freeloaders flourish' (idem: 200). Human nature is fragile and vulnerable because it is rooted in a complex subjectivity that makes persons so different from one another. There always has been and always will be 'too-altruist' persons and cold-hearted persons, but is should never be up to other people to decide on their humanity on basis of their ability to fit into a human (all too human) account of what is human. If the values with which we characterize human action do not reflect the subjective attitudes, feelings, and thoughts of human personhood, then they loose what ever could make them human in the first place.

Faulkner writes in his *As I Lay Dying* about the indestructible subjectivity in every person, in this case the queer nature of Darl who, albeit his profound 'natural affection' for other human beings, continues to defy the normal reasoning of the Deep South family: 'But I ain't so sho that ere a man has right to say what is crazy and what ain't. It is like there was a fellow in every man that's done a-past the sanity or the insanity, that watches the sane and the insane doings of that man with the same horror and astonisment' (Faulkner 1996: 226).

Chapter Two

Personhood in Nature and Culture

Moral facts and values exist and are personal. We feel them, we act upon them, and we disagree about them. This is what intuitive, Humean backgammon-playing experience tells us. But, as Kant once observed: 'There is something splendid about innocence; but what is bad about it, in turn, is that it cannot protect itself very well and is easily seduced' (Kant 1998a: 17/404-5). We live in a world that does not know anything about our ethical doubt, convictions, and conflicts. So the job of philosophy must be to try to understand how moral values are facts and how they fit in with the rest of what we know about nature. This is not just another futile theoretical investigation of little interest to others than academic philosophers. On the contrary, it has important personal, social, and even political implications, since our way to think about value and ethics influences our decisions, attitudes, and behavior with regard to the other, both those we care about and disagree with and those that we do not care about and still disagree with.

In the foregoing pages I have repeatedly said that I am naturalist. Now, by way of conclusion, I want to clarify what I exactly mean by this. Further, I have promised to explain how the relation between ethics and subjectivity might shed some light on the process of naturalization. I approach these two issues in three moves that together, hopefully, will clarify my position. First, I specify what I mean by naturalism by positioning it with regard to other kinds of naturalism. Then I question the intuitive relation between subjectivity and moral relativism and argue that the relation between subjectivity and ethics is not a plea for moral relativism. And finally, I conclude with a brief summary of some central arguments of the dissertation about the natural status of ethical personhood in a non-ethical world, and which is, at the same time, a modest suggestion for further process of naturalization.

This concluding clarification is, as mentioned earlier, obviously not an attempt to settle the long-standing and very complicate debate about the fact-value distinction. This exceeds both my capacity and the scope and aim of this work. The work is meant as an argument for the irreducible relation of subjectivity and ethics with respect to a naturalized conception of human personhood, and I believe that this needs a final discussion of both naturalism and relativism. The previous chapter criticized neuroethics for excluding subjectivity in its attempt to naturalize ethics. This chapter can be seen as an extension of that argument by showing that the notion of subjectivity does not block the process of naturalization nor lead to moral relativism, which is a common objection to the notion of subjectivity in ethics. On the contrary, a clarified notion of subjectivity can be seen as a part of this

process, because it can help to give a more solid account of a naturalized ethics than is proposed by the neuroethical exclusion of subjectivity.

Which Sort of Naturalism?

My criticism of neuroethics stressed the notion of subjectivity and tried to argue for the implausible and sometimes even dangerous solutions offered by an ethics controlled by impersonal functionalism. Subjectivity matters, or rather, subjectivity is fundamental to ethical reasoning. When an ethics ignores the individual nature of the person or 'the specific contingencies of one's creaturely point of view' (Nagel 1986: 9), i.e., subjectivity, and wants to explain what the person should do and think with reference to neurophysiological mechanisms and a mythical transparent fit-functional society that are both incomprehensible and far from transparent for most persons (if not all), then it stops being an ethical theory and transforms into bad moralizing. Many neuroethicists (and perhaps even more scientistic philosophers) abhor religion and talk about the soul because such accounts are grounded in an obscure metaphysics that ignores the actual state of the universe and the biological being of the human person. Sometimes, however, they adopt a patronizing tone explaining that religious beliefs are just as natural as anything else in the universe, because the brain needs to form religious beliefs just as it needs a panic center to avoid danger. This is all due to the simple fact that '[h]umans are belief-formation machines' (Gazzaniga 2006a: 161). In fact, many creative persons (e.g. Socrates, Dostoevsky, Flaubert, and Van Gogh) and religious leaders (Moses, Buddha, Muhammad, and Paul) may very well all have been suffering from temporal lope epilepsy, which caused them to form their art and religious systems (idem: 156-60). But often religion and metaphysics are simply considered as dangerous fiction. The danger is that such accounts can inspire us to act on false beliefs about the world and human nature. This is possibly true, but there is something crafty about this kind of reasoning, since the truths proposed by neuroethics and scientistic philosophers often contain the same sort of obscure dogmatism that they accuse religion and metaphysics for. Evolutionary theory comes in all kinds of flavor, explanations for physicalistic determinism in even more, brain imaging techniques still only penetrate the upper-layers of the brain lopes, and the complexity of the brain-mind-body unity that we call thinking is today not more transparent than that of the universe. And the constant pledge that in a (near or distant) future we will eventually discover this or that which can finally solve traditional philosophical problems sounds a lot like the future-directed explanations of some religions (wait and see what shall come!).

This is, of course, putting things on the edge. The vast majority of serious empirical scientists is of course very cautious with their conclusion about human nature and remains within their respective (local) fields of investigation. However, the tension between natural sciences and the humanities can be strong, and I think that this is due to an implicit dichotomy in the conception of nature. A dichotomy that has existed since the birth of modern science but recently grew into an intense public debate.

The last decade of the 20th century was tormented by the so-called 'science wars' between different academic cultures. On the one side, we had the 'postmodernists' who, backed up by Thomas Kuhn's famous book from 1962, argued that 'scientific theories are purely social constructions with no more claim to the label "truth" than other socially constructed ideas' (Baringer 2001: 9). On the other, the scientists who retorted that the scientific method is the only way to avoid mysticism and personal biases which will divert us from logic and rationalism and put us on the road to dogmatism (idem: 5-7; for a good discussion of the science wars, see Ashman and Baringer 2001). This war is still fought today, although postmodernism is substituted with religion. The public and academic discussion is no longer about postmodernism and scientism, but more about how scientific progress and (Christian) religion are the two pillars of western civilization. The question about naturalization is too often seen as a choice between either the cold world of science or the colorful world of religion. A naturalist has to confess complete adherence to the impersonal account of science otherwise he or she is considered as logically inconsistent or simply superficial. On the other hand, a religious person must account for the paradox of creation and afterlife if he or she is not to be labeled a conformist and intellectually lazy. And although talk about religion may seem irrelevant for the present discussion about subjectivity, ethics and naturalization, I think that the fear of being called mystic, idealistic, or even religious affects the way much discussion about nature is conducted. I will go so far as to say that one of the most difficult issues in the process of naturalization is an implicit dichotomy in the conception of nature: *either* objective, certain, and realistic *or* subjective, uncertain, and idealistic.

This has recently been discussed by Hans Fink in a thorough and clear article about ethical naturalism and the concept of nature (2006). In a discussion of McDowell's famous distinction between two sorts of naturalism, 'neo-Humean naturalism' or 'bald naturalism' (nature as revealed by the natural sciences) and McDowell's own 'acceptable naturalism' or 'Greek naturalism' (nature understood in a broader sense that include all human potentials, e.g. virtues, skills etc.), Fink introduces a third sort of naturalism that he calls 'unrestricted or absolute naturalism' (idem: 205).

While being in general agreement with McDowell's view on a more acceptable ethical naturalism, Fink believes that the third sort of naturalism is better, because it can 'point us in the right direction when it comes to the understanding of a positive naturalism providing a convincing alternative to bald naturalism' (idem). Here I will not deal with his comments on McDowell, but focus on his arguments about the underlying conception of nature involved in the different understandings of naturalism.

Fink points to the ambiguity in our use of the concept of nature. On the one side, we talk about 'the nature of something – e.g. my nature, human nature, the nature of a certain mineral, something being in the nature of things' and, on the other, we talk about 'nature as a realm of its own, the world of nature' (idem: 206). The first way of talking about nature must include absolutely everything, and '[i]t would be wrong to leave anything out' (idem: 207), since every aspect of the being x is part of the nature of x. All that we eventually can come to know about the nature of a certain being is releveant to the nature of that being. Say, for example, that some day an attentive scientist discovers a butterfly that can actually swim under water, then this discovery will be very important to our understanding of the nature of the Lepidoptera order of insects. Fink calls this a '*non-contrasting* conception of the nature of x open to any future additional information about x' (idem: 208). In contrast to such an unrestricted understanding of nature, we have the other conception that wants to understand the nature of the world we live in. Since we want to find the basic and constant structure of the world that we try to understand, such a general conception quickly leads to contrasts and restrictions that 'would then identify nature with certain parts or sides of the world' (idem: 209).

This is, I would claim, what happens in the neuroethical account of values. Moral facts are functional facts, because the world is functional in nature. The body is a complex functional machine, an autumn leaf in the wind is just as mechanical as the words of Shakespeare and my kitchen blender; the society has to reflect this functionality and the person simply fit into the machinery. In this way we can solve the tedious issues of moral disagreement and prevent that the by-products of evolution (an unfunctional moral feeling like 'pure altruism') come in the way of the progress of evolution, otherwise this insatiate progress will sweep us away like insignificant crumbs on a table. Neuroethics comes to this conclusion on the background of a restricted and fragmentary conception of nature. Since the physiological constitution of human nature and the general laws of physics and biology are functional, the basic structures and dynamics of nature have to be

functional, and explanations of nature must conform to these basic features if they are to be considered valid and true to *real* nature.

Fink's argument for an unrestricted or absolute conception of nature contests such a restricted view on what is real nature. We cannot operate with a simplistic conception of what is nature and what is not, because '[n]ature is never *mere* nature. That which is *more* than *mere* is nature, too [...] the fact that there are so many incompatible contrasting and restricted conceptions of nature and the fact that each of them creates formidable difficulties in accounting for that which is regarded as non-natural provides a strong motivation for trying to explicate an absolute conception.' (idem: 217). Fink is well aware that such a conception does not solve the problems of naturalism, but it is a way to avoid 'the otherwise endless oscillation' between the different conceptions of nature: 'Nature is that which all possible domains are domains of. Nature is all there is, all that is the fact, all that happens' (idem: 218). Hence, when I define myself as a naturalist, it is this sort of naturalism that I refer to.

However, I need to be more specific if I want to position myself properly with regard to the process of naturalization. My primary interest here is to clarify what I mean be naturalism with regard to ethics and subjectivity. This, I will claim in the conclusive section, bears a consequence for the more general conception of naturalization.

In Fink's absolute or unrestricted conception of nature there is no contradiction in speaking of the nature of God or of the super-natural (idem: 206, 218). This is certainly true on a theoretical level. The Christian God, Odin, Demeter, E.T., and golden mountains all have a nature and therefore belong to an unrestricted conception of nature. They tell us something about nature in general, as the sum of what *is* actually experienced, without disqualifying some experiences of nature on a theoretical level.

On a practical level, things are a bit more difficult, since the different experiences of nature often disagree and conflict. And an important aspect of my conception of ethical reasoning is that such disagreements must remain on a practical level, and not be settled with speculative or ideological arguments. With speculative and ideological I mean ungrounded arguments and blind convictions that presuppose or impel more than can be accounted for with regard to common human concerns. This, however, does not imply that we can do ethics without theoretical arguments. But such arguments are different from speculations and ideologies, since they are (or should be) impartial and rooted firmly in the nature of human practices (our care, interest, and concerns). This distinction between speculative/ideological and theoretical arguments is important in order to

understand my conception of ethical naturalism. This means, for example, that our conception of God or other super-natural natures must not a priori restrict other conceptions of nature. In the naturalistic perspective that I propose it would be wrong, i.e., bad practical reasoning, to say that it is the will of the Christian God that missionaries remove preservatives from communities in Kenya because it is a sin to have protected sex. Such practical reasoning disregards the fact that commonly people do want to have sex without thinking about children and even less about dying from a terrifying virus. The same argument holds for neuroethics and other restrictive kinds of naturalizations. We cannot use one conception of nature, in this case a functional-evolutionary account, to determine how we should think about ethical questions that involve the concerns of persons who consider themselves to be more than just objects of impersonal functions. Conceptions of nature that a priori exclude religious phenomena or what, in their conception of nature, is considered to be unnatural phenomena such as personhood, pure altruism, or protected sex are making *one* aspect of nature into an *absolute* conception of nature. Fink's third sort of naturalism argues that such a move is implausible and restrictive because it wants to explain the whole by isolating a part and contrasting it to the whole. Instead of using specific scientific methods and insights to define a restricted conception of nature, Fink proposes that 'unrestricted naturalism takes it to be defining of science that all things are the measure of science' (idem: 218).

Therefore, my conception of ethical naturalism claims that subjective experience matters, because the ethical qualities that the subject experiences in its interaction with the world (goodness, badness, love, shame, resentment, wrongness, rightness, courage, virtue, commitment, etc.) are a natural part of what it means to be a human subject, a person. The deflated ethical theory that I propose assures the existence of such thick moral concepts, since they are a substantial part of human experience, but argue that they can be examined and discussed without framing a specific moral space severed from the rest of what we know about human nature. Our thick personal concepts that regard religious experiences, fragile identity, vulnerable personhood, esthetic experience *and* the cold insights from the empirical sciences such as our mammal kinships, evolutionary past, neurophysiological structures and mechanisms, social behavior and statistics, economic and political dimensions are, together with an infinity of others, all part of human nature. And to neglect or intentionally disregard any of these aspects of human nature is to impair ethical reasoning. And this is what I believe happens if our conception of nature is infected with the implicit dichotomy that haunts much ethics (and philosophy in general) today.

As I mentioned in the beginning of this section, I think that this dichotomy (objective, certain, realistic vs. subjective, uncertain, ideological) has been stressed by the turmoil of the 'science wars' of the last century. Subjective experience becomes associated with introspective, unscientific idealism with a twist of mysticism, which has nothing to do with real solid science. It becomes an either/or, and not a both-and. We can have a subjective taste of wine and a subjective feeling of art, but not a subjective feeling about ethics, because then we end up with moral relativism that does not help us in our existence with other people. On the other hand, if we consult empirical sciences such as neuroscience and biology then we necessarily end up with a cold and impersonal account of our warm ethical relation to other people.

Such ways of positioning oneself to the dichotomy can be seen in the writings of Taylor, Blackburn, and Ricoeur. Although all three thinkers have worked much with the natural status of human personhood and thereby the dichotomy itself, their work somehow reflects the excluding tendency that affects much writing about naturalism. Taylor's constant rejection of naturalism tout court and of deflated naturalistic theories of ethics such as Blackburn's quasi-realist expressivism reveals a general conception of naturalism as impersonalizing the subject and stripping it of all that which makes it a human subject. And although Blackburn's account of practical reasoning, which I follow, places subjective experiences in the center of the theory, he does not consider religion or religious experience to have anything to do with ethics. I believe both these attitudes to be wrong, since they neglect two rich sources of information about human nature. And although I have not considered religion in this work, since my primary focus was on the natural relation between ethics and subjectivity, I believe it to be an indispensable aspect of any ethical theory, since religious experience reveal something about human nature, for example, the experience of transcendent alterity in subjective experience. Ricoeur's writings, although more open-structured, suffers from the same dichotomy. His philosophical and theological analyses are strangely segregated, and his discussion of naturalization is only peripheral. Nonetheless, I think that Ricoeur's theory of subjectivity, because of its deep and open-structured analyses, offers, when reformulated, an adequate framework for an unrestricted conception of nature and a naturalistic approach to ethics and subjectivity in the sense described above.

Thus, said roughly, our conception of what is nature and what is not affects the ontology in which we ground our ethical practices and argue for our ethical values. The ontology that I have been argued for is constituted by care. This is, of course, a human way of talking about it, because it is hardly likely that a tick cares about its existence in the same way that we humans do. However,

every living organism 'cares' about living in the sense that it interacts with the environment in order to live and reproduce life, from the functional splitting of the cell to human care for the other in erotic and parental love. Human ethical values are rooted in and generated by this basic care that characterizes all living organisms, although human care is not governed by the same functionality that qualify other mammalian care. To understand this natural conception of human care and not reduce it to what it is not, we have to make our conception of nature as broad as possible so that we do not start on a wrong ontological footing. At the same time, it is important to argue why such a conception of a naturalistic ontology is not the same as to say that everything goes, and that we cannot exclude anything from our conception of human nature. As I said, ethics is a practical matter, and as everybody knows that not all practices are good practices, and that some are better than others. Therefore, to argue for an unrestricted conception of nature in ethical reasoning is not to argue for moral relativism. The relation of subjectivity and ethics has something to say about what is like to be a human person.

So, before concluding with a final summary of the *natural* relation between subjectivity and ethics and why this relation is important to the general process of naturalization, we have to look at one last danger when talking about a naturalistic conception of the relation between subjectivity and ethics, namely the accusation of opening a door for relativism. I will defend the ethical relevance of notion of subjectivity by criticizing very briefly some methodological aspects of a recent attempt to argue for relativism on the background of a extensive account of emotions, subjectivity, and naturalization.

Not a Plea for Relativism

Subjectivity must not be confused with subjectivism. Subjectivism has a bad reputation in moral philosophy and is often associated with relativism, idealism, and egoism. If ethics were to depend on the values of the individual subject, how could we then possibly avoid that everybody acts on what that person thinks or feels is best for him or her? Or even worse, how could we prescribe or demand that the person did otherwise? Questions like these trouble every attempt to put the subject at the center of moral reasoning, and that with good reason. Few philosophers would recommend strong moral relativism, since the sense of doing moral philosophy is to find the best way for different persons to interact with one another. And relativism makes this difficult. If I cannot say to another person that he or she has treated me wrong, insulted me or hurt me, human intercourse becomes difficult, if not to say impossible. Human lives are interwoven, and people care about one

another. This is a fact that even the most subtle speculation cannot dissolve. Why, then, would anybody recommend something as nonsensical as relativism? There are, however, good reasons for this. As we saw above, morality can turn ugly and become bad moralizing that does not respect the individual personhood of every human being. Moralizing is different from morality because it relies on an authority that claims to know what is best for a person or a community without regarding what the person or community think themselves. History and present day are full of examples of bad moralizing and of people saying what other people should think, feel, or do. I am making a normative claim when I say that some neuroethicists do bad moralizing, or that some religions and communities do bad moralizing systematically, make it a policy, and cause dismay and terror around them. Such a claim can easily turn into bad moralizing itself, if it is not supported by good arguments. The question of what is a good argument and what is not is also a normative claim. So, as I have tried to argue for earlier, there no way of getting behind normativity. We are tethered to normativity in virtue of being a human subject who lives as a person among other persons. And, on this view, ethical theory is basically the attempt to argue for what is considered a good argument or reason when judging, acting, and arguing morally. In short, how do we legitimize the values and norms by which we interact with other people?

The picture that I have laid out in the foregoing analyses focuses on subjectivity and, in particular, on feelings as the source of normativity. And I have argued, from within the framework of Ricoeur's theory of subjectivity, that this does not qualify as relativism. On the contrary, relativism is not an option, because self and subjectivity can not be reduced to egoism or skepticism with regard to the values of the other. My subjective feelings are not constituted by my own limited understanding and experience of the world or restricted to a concern for my own well-being. I have argued for this in terms of the practical space of reason and an ontology of care. My basic claim is that values are constituted by subjective feelings, and I have argued that these feelings cannot be understood without considering that the subject is a person other than being a biological being. But before concluding with a short outline of how I see a successful naturalization of personhood, I will confront the problem of relativism by presenting some aspects of a recent philosophical argument for moral relativism that refers to notions that I myself have been stressing, namely human concerns and subjectivity.

When discussing different theories of emotions in the last part, I briefly mentioned a contemporary variant of the James-Lange feeling theory proposed by the Jesse J. Prinz (page 121, note 24). Prinz is very productive and has, since 2002, published three books that he considers to be one extensive

argument for a Humean empiricism, which parallels the structure of the three books of Hume's *Treatise*: concepts, emotions, and morality. Modestly, he characterizes his books as 'footnotes' and yet 'a tribute and modest extension of Hume's masterwork' (Prinz 2007: vii). Prinz' work has an impressive interdisciplinary scope, comprising cognitive science, cultural anthropology, neuroscience, and evolutionary psychology, and his provocative and clear style makes his books exciting and accessible for readers coming from many different fields. From my own point of view his position is interesting, because he defends a feeling theory in emotional research and, further, sees emotions as essential to ethical theory. Thus, my own framework is similar to his, and yet I completely disagree with most of what he has to say about both emotions and ethics. I think that this can be explained by our contrasting understanding of the notion of subjectivity. This difference makes his theory of emotions and his ethical theory utterly different from what I have proposed in this dissertation. Obviously, I cannot give Prinz' elaborate account the extensive treatment that it merits, but will only deal briefly with some methodological aspects that separate his approach from mine and makes his final result, moral relativism, a *logical* conclusion that I disagree with.

Prinz has an apparently clear agenda. Empiricism can answer most philosophical questions[66]. Therefore a philosophical theory must work conceptually with what we learn from biology and cultural studies. For example, the study of emotions should not focus on one of the two, but realize that '[a]dult emotions are both cultural and biological. One needs a notion of compatibilism to capture this interaction' (Prinz 2004: 158). And he argues for the same 'biocultural approach to morality' (2007: 270), because 'moral rules are not innate. Rather, they emerge through interaction between biology and culture […] Biologically based behaviors are not quite a constraint on the genealogy of moral rules, because culture can override them. We can understand many moral norms as culturally specific variations on the same biological themes' (idem: 274). However, he wants to reject biological reductionism, because the innate gut reactions (sadness, fear, surprise, panic, disgust etc.) encoded in our genes is overridden by culture that 'does not merely activate innate programs; it rewires our moral software […] One might put the point by saying that humans instinctively take their clues from culture. Culture also converts our biological norms into moral norms – norms that are grounded in sentiments and extended to third parties' (idem: 286). This clarification gives him an empirical framework wherein he employs the conceptual analyses that were prepared by his first work *Furnishing the Mind: Concepts and their Perceptual Basis* (2002).

[66] He writes, for example: 'I think that enduring philosophical questions can be illuminated by empirical results, and, indeed, they might not endure so long if we use the resources of science' (2007: 9).

Here I will focus on his understanding of the generation and status of human values on the background of this biocultural framework.

Whereas a functional understanding of values induced neuroethics to suppress subjectivity, Prinz cares very much about defending the individual subject's right to 'privacy, free expression, and diversity', even to an extend that secures 'isolationism' (idem: 212). In this sense, his agenda is the complete opposite of Casebeer's version of neuroethics. And still, there are many similarities in the ways they construct their notion of values. We have two sources of information about human values, biology and culture. And as a good empiricist, Prinz does not accept other foundations of values. This is basically the same reasoning that we found in neuroethics only that Prinz' naturalistic approach starts with a different assumption. Contrary to neuroethical functionalism that coins our values on the workings of biology and *society*, Prinz sees values as constructed by biology and *individuality*. We cannot assume that moral values possess the stability that neuroethics or other ethical theories presume them to hold. Nonetheless, 'moral properties are perfectly real', since 'wrongness induces emotions in us, and those emotions impact our behavior' (2007: 167), but still '[m]oral properties are not objective in any interesting sense' (idem). Moral values are facts, since they are 'causally efficacious' and bring about observable changes in the environment. But they are different from physical facts, because they depend on the subjects who deal with them. A mountain does not disappear just because we want it to, nor will the sun pierce through the sky on a rainy day, because I desparately need some light. These are facts that exist in another way than moral values, because they are absolutely independent of the constitution and will of the subject. The factuality of morality and norms are different, because '[m]oral facts are like money. They are social facts that obtain in virtue of our current dispositions and practices. They are as real as monetary values and even more important, perhaps, in guiding our lives' (idem). So, if values are facts but still not objective facts, then what kind of moral theory does Prinz argue for? He calls his position 'constructive sentimentalism', since 'the members of different cultures can have different fundamental values' (195). Values are constructed by the culture that they are supposed govern.

Prinz conjoins the sentimentalist tradition in saying that subjective feelings are fundamental for ethical thought. In fact, 'moral values are sentiments' (120), which means that they are expressed by the emotions of the subject who judges something as wrong or right. So far he follows the kind of reasoning that I have been arguing for. Sentiments represent our basic concerns, i.e., human dispositions to experience different emotions (84). And he gives many examples of moral

sentiments (i.e. moral values) such as anger, contempt, resentment, fear, sadness, happiness, liking, disliking, loving, hating, and so on.

This all fits well with my own account of human values. We cannot neglect the subjective aspect of values, and we saw that also Blackburn used the metaphor of monetary values to express that we cannot find moral values but in the way that human subjects interact with one another. But there the similarities stop. Prinz' has a strict empirical framework that he must respect. Whereas Blackburn and the sentimentalist tradition that I adhere to retain that moral values cannot be accounted for with truth-conditionals and reference to 'hard' empirical facts, Prinz sees it otherwise. He is convinced that we have to stick to what we know. If morality shall not be completely unnatural in a perfectly natural world, then we need something to hold on to when we discuss moral disagreements and practical issues about how human beings should judge one another's behavior. Prinz therefore spends a lot of time criticizing the view that I have been defending, namely that human values does not depend on truth-conditionals in the same way that functional logic does (e.g. 100). On the contrary, he claims that moral values do indeed 'refer to real facts' (108). We have to be able to point out empirically why one action is wrong, while another is right, otherwise it 'would be hard to make sense of moral discourse' (idem). We cannot retort that practical human reasoning is not a matter of truth in the sense that a realist semantics accounts for truth (i.e. by reference to empirical facts and logical deduction), since this goes against the naturalism that Prinz defends, i.e., the facticity of empirical findings (from 'neuroscience, psychology, psychiatry, anthropology, cultural history, and ethology) that make 'philosophy as continuous with science' (10).

As I mentioned earlier, Prinz is interesting because his empirical framework reaches the complete opposite result of Casebeer's functionalism, although they work with the same confidence in empirical sciences and logic (biology, culture/society, and conceptual analysis). Prinz argues for a strong relativism that secures the right of the value system of the individual agent in front of other agents (Prinz often uses the term agent instead of person as if to emphasize the logical nature of his analyses). For example, the evaluative propositions such as 'killing is wrong' is both content and truth relative in the sense that the proposition depends on the individual agent who utters it and is not necessarily wrong for the agent who is in fact killing another human being (175-83). This seems to be the logical implication of adopting an ethical theory that puts the individual subject and his or hers feelings at the center of moral values. How can I say that another person's values are wrong if I claim that human values are constituted by subjective feelings? Prinz thinks that his argument for relativism is evident, because it is backed by the empirical fact that different persons do feel

differently in front of the same thing or event, and thus 'the existence of differences in people's sentiments entails a difference in moral facts' (175). Therefore, he uses many pages to argue for the implausible justification of the common western conviction that cannibalism and female circumcision is categorically wrong. Obviously, he does not appraise such practices but questions our firm belief that nobody wants to be eaten or have their genitals mutilated. He observes that 'cannibalism may be the default human practice' (224), and he then refers to many studies from cultural anthropology that confirm this strange *fact*. With regard to female circumcision, he argues that many girls are willingly accepting this ritual, and '[a]s far as I know, most women who have come to accept circumcision have not undergone any unusually coercive methods of indoctrination' (210). We need understand that '[r]elativism promotes tolerance in cases where the "victims" are truly willing' (211).

However, I believe that the evidence of Prinz' arguments for moral relativism that depends on the theory of subjectivity that one adopts. And I think that his evidence becomes less evident if we consider how he accounts for the 'subjective' nature of human care and concerns: 'Emotions, as I have argued, represent concerns. Concerns are organism-environment relations that bear on well-being' (85). Our concerns are constituted by the two basic reference points that make up Prinz' empirical framework: biology and empirical environment. These are the sources of the firm facts that we need to refer to if we want to understand human values. The only problem is that the individual subject that Prinz cares so much about does not have much to say in this account, because he or she is characterized by what we know about the biology and the environment in which the subject lives. Our values are subjective and real, but the notion of subjectivity remains a somewhat frozen object of 'angry' emotional forces that reflect the mammalian 'altruistic' behavior that is never really moral, but an expression of a genetically encoded seeking of pleasure and proper well-being in a specific environment. Thus, Prinz does not think that 'concern is the most fundamental moral emotion; the emotions that figure in moral judgments express blame for the wrongdoer, not sympathy for the victim' (98). And this because he is certain that 'there are two fundamental classes of other-directed moral emotions (moral anger and moral disgust) and one derived class (moral contempt)' (76). He supports such convictions on his empirical study of emotions, where human emotions are seen as 'natural extensions of basic emotions that arise in the contexts of transgression […] Rights violations are threats to persons. Sexual mores involve threats to the natural order. Rank violations are threats to the natural order of persons' (idem). Therefore, when Prinz argues that he performs a descriptive analysis of normativity in contrast to the

'imperialistic ambition' of Kant, Mill, and Aristotle whose normative agenda ignores the fact that 'every rule has its own biocultural history' (305), he completly forgets that his theory of biocultural sentiment grounds in a universal and very normative claim: the human subject is controlled by basic emotions that together with factual environmental input creates a human person whose concerns and cares are just variations (meta-emotions and meta-sentiments) of a more fundamental biological urge (gut reactions) for homeostatic well-being.

If we go back to Prinz' argument for the reduction of the phenomenological variety of human emotions to gut reactions based on 'core relational themes' that 'are directly relevant to our need and interest' (2004: 66; jf. pp. 121 note 24), we can see that he does effectuate a reduction, although he claims not to. Prinz tries to camouflage this covert biological reductionism by arguing for the limits of evolutionary ethics and the priority of the individual moral sentiment, but it seems somewhat crafty when he has a very clear idea of the biological origin of such sentiments. The subject is free to act according to its own individual morality, but this freedom is constructed by less free feelings about what it should care about. One might be tempted to say that the individuality that Prinz argues for is rooted in his conviction that every subject is biologically wired to care primarily for his or hers own well-being.

Now, I disagree with this naturalized picture of human feelings of concerns and care, because I find his empirical account of subjectivity reductive with respect to what I consider the (f)actual subjective experience of feelings of care and concern. To say that human care and concern originate in basic mammalian affect programs (which I agreed to in the last chapter) does not evidently lead to the conclusion that human values and human care and concern are restricted to the basic emotional reactions that we find in other mammals. If we use the embodied (biological) nature of our emotional experience to disqualify the subjective nature of such experiences, then we lose what makes these experiences particularly human[67].

I simply think that Prinz misses something fundamental in his 'compatibilist' attempt to account for the biological and cultural constitution of human subjectivity. Human nature is not constituted by nature and culture, but is a person who is embedded in both nature and culture. To deduce human values from a conceptual analysis of biological and cultural norms is to leave the subject out of the picture. If everyone could remain within their own isolated system of values, we would not

[67] Goldie points to the same problematic reduction if his otherwise favorable review of Prinz' book on emotions: 'Whatever one's view of the expressive properties of music, it does seem a distortion of our notion of what makes sadness appropriate to say that we are 'in error' when we feel sad listening to Beethoven's late string quartets. This is just an example of the kinds of difficulties faced by accounts that reduce the appropriateness of an emotional experience to its being in accord with what we are 'set up' or designed for (Goldie 2006: 457).

need an ethical theory, or an idea of personhood for that sake. Apart from my personal doubt that anyone actually likes to be eaten or mutilated "willingly", then the fact remains that subjects cares about being a person who cares about other persons, although these do not share the same values. Just because moral relativism is a logical possibility considering some accounts of the biological constitution and cultural diversity of human beings, this does not make such a theory either more true (descriptively) or attractive (normatively). As Blackburn writes: 'An immature, unimaginative, unsympathetic, and uncultivated ethic might be quite coherent, in the way that the Decalogue is quite coherent. But the people who embody the attitudes it commends will not be particularly admirable' (1998: 310). An attempt to naturalize personhood can not fall for the functionalist or relativist temptation, if it wants to account for the fact that the subject cares about being an individual person *and* a person who lives together with persons with different values.

I will now conclude with a short outline of my basic arguments and a modest suggestion to we can naturalize human personhood without losing the complex relation between subjectivity and ethics.

Conclusion: Naturally the Subject is a Person

It might appear somewhat unorthodox to use Ricoeur in relation to the question of naturalizing personhood. I think, however, that Ricoeur's theory of subjectivity has something important to say about the process of naturalization. In particular, the irreducible relation between subjectivity and ethics that was emphasized in my reformulation of his theory is relevant for the contemporary debate about the natural status of the human subject. I believe that it is so, because naturalization often seems to want to get rid of the subject or reduce ethics to something that it is not. Ricoeur analyses show that the reality of human values is multilayered and constituted by heterogeneous aspects of human life such as self, body, historicity, world, practices, and the other subject. I believe that this theoretic core of his theory is first fully disclosed in a reformulation that focuses on the complex notion of subjectivity. I have much confidence in the importance of the peculiar nature of subjective experience. I am convinced that our conception of nature and, consequently, of ethics is incomplete if we do not clarify the experiential dimension of the nature in and through which human beings live. At the same time, Ricoeur's analyses are precious because of the strong emphasis on the alterity in human experience and being. The notion of subjectivity becomes incomprehensible without an examination of that which is not the self: body, world, other.

In this part I have tried to clarify my basic argument for the irreducible relation between subjectivity and ethics by showing that a naturalization of ethics necessarily involves the question

of human subjectivity, if it is not to distort the notion of human values. And part two argued that the question of subjectivity cannot be separated from the notion of personhood, because the human subject is characterized by feelings that reveal a basic care for being a person among other persons. I have thus argued that the notions of subjectivity, ethics, and personhood are irreducibly related to one another. We cannot understand one of the concepts in isolation from the other two. This is due to the particular ontological status of the human subject.

In order to understand subjective experience and action, we have to formulate an ontology of care that respects the fragility of human existence in a world that does not care about the fact that the subject cares about the world and other persons. Human experience is affected by this fragility is the sense that the subject experiences the world and the other persons as both different from and, at the same time, part of its own being. The care of the subjective person affects the experience of that person. Subjective experience of the world is not an indifferent registration of things, events, and persons, but configured by the heterogeneous values that the subject harbors in the interaction with the experienced things, events, and persons. This interactive configuration of subjective experience is what I have tried to capture with the notion of the practical space of reason. Human experience is primarily practical and enveloped in the affective dimension of these practices. Care, interest, concern, desire, fear, insecurity, panic, joy, hopelessness, absurdity, love, hate, hope, and so forth are all affective experiences that arise in this interaction between the self and that which is not the self (world, body, other person).

This relation between selfhood and *alterity* in human subjectivity is the core of the framework that I have constructed by my reformulation of Ricoeur's theory of subjectivity. Ricoeur's theory is a life-long development of this relation between the voluntary and involuntary aspects of the human subject in terms of spontaneity and receptivity, reason and body, presence and distance, immediacy and historicity, self and other. The fragile affectivity in the heart of the subject is characterized by this structural non-coincidence in the being of the subject. The subject does not exist but in that which is not self, and subjective existence is a constant affirmation of selfhood through this alterity. Selfhood, however, cannot be affirmed in the bare existence of the self, but needs the presence of other subjects, other selves, that recognize the personhood of the self. This is due to the fact that the world of the subject is constituted by the coexistence with other subjects. The alterity that the subject encounters in the heart of its being is therefore primarily constituted by the *active* alterity of other selves. When the subject affirms itself in the alterity of its existence, it does so through the presence of other subjects who experience, evaluate, and suffer these actions. The actions of the

subject express the personhood that the subjects care about, because these actions inscribe themselves in a common world shared together with other persons, a human world.

We cannot abstract from this humanity without losing sight of what makes the subject human. Human thoughts, words, actions, and practices are all characterized by the personal qualification that differentiates these from non-human actions and mere anonymous events of impersonal nature. The normative dimension of subjectivity is rooted in this particular nature of the subject. The subject is a person by the simple fact that it exists, and yet this fact is a question of how the subject relates itself to the fact that it is a person who depends on the recognition of other persons. Thus, Ricoeur says that to be a person is a task. The subject cares about being a (certain) person, and this care makes the subject see 'itself as another', because personhood is primarily constituted by the presence of other persons.

There are however more to the being of the subject than self and the alterity other persons. The subject is embedded in and constituted by the alterity of the natural world whose laws and rules do not respect personhood or human values. We cannot understand the being of the subject without considering this aspect of human nature. We experience this alterity in different degrees, through our embodiment (perspective, physical constitution, limited actions) and through the absolute passivity in front of cosmological time and eventually death. We are born, we age, and we die.

The reformulation of Ricoeur stressed these two fundamental aspects of human experience of alterity. We are persons *and* biological beings. Human subjects are constituted by the impersonal workings of the natural world *and* by the personal nature of humanity. This is what I have argue for in this dissertation by examining the ethical and biological being of subjectivity. Ricoeur's theory remains the theoretical framework for my further analyses, although I have developed his arguments in directions that might seem distant from his. This, however, is one of the virtues of his theory. His arguments and conclusions remain open-structured without being superficial. He is clear about what he thinks, but his analyses invite to further reflection. I have done so by examining the *natural* relation between subjectivity and ethics.

I have spent much time on arguing for the natural status of human values. Human experience is affective experience. We experience the world as qualified by our feelings. And these feelings reveal the values by which we experience, think, and act. This is why I have insisted that subjective experience is configured as a practical space of reason that is structured by heterogeneous values generated and refined by the world, other, and the self. Human values cannot be examined in isolation from the subjective experience of such values. The natural status of human values does not

suffer from this subjective aspect of their constitution, because this subjectivity is not a disqualification of their status as being part of nature. Human values are natural facts about human nature just as the constitution of human bones, railroads, and cell phones. They may be more fragile and difficult to handle, but their existence is factual in the sense that they affect the world and the beings in that world in a very concrete manner. Our experience of values plays an active role in the constitution of the world in which we and other creatures live. Therefore, to disregard or abstract from this experience of values is to transform the values that we all want to understand.

This was the central point in my critique of both neuroethics and Prinz' version of moral relativism. In their attempt to make ethical values less difficult to handle, they made them into what they are not, functional and strongly relativistic. Neuroethics neglected the fact that human persons are more than the society in which they live. Human persons are individuals with all the blessings and the problems that this implies. Prinz completely misses the other fact that individual subject are also persons. The subject's values express more than the individual nature of the subject (desires, preferences, inclinations, upbringing etc.). Values express the humanity in which this subject exists as a person among other persons. The subject cares about being more than an individual, or as Strawson puts it, we are individuals because we are part of the common roots of human personhood. The individual differences of human values are only possible on a common agreement of what it is to be a human person. Values vary from person to person, and even more drastically from culture to culture, but they are still human values that characterize human behavior (i.e. care, love, hate, interest, and concerns). The job of a descriptive ethical theory is not to make the different expressions of the values incomprehensible to one another, but to make these capable of interacting in ways that do not compromise the coexistence of the human persons who have different values. This is what I mean by the natural existence of moral values. And an understanding of human nature is impaired if it do not respect that the factual existence of moral attitudes, feelings, and behavior is irreducible to anything other than moral values. As Iris Murdoch writes: 'The authority of morals is the authority of truth, that is of reality. We can see the length, the extension, of these concepts as patient attention transforms accuracy without interval into just discernment. Here too we can see it as natural to the particular kind of creatures that we are that love should be inseparable from justice, and clear vision from respect for the real' (Murdoch 1967: 90-1).

Therefore, I consider human values to be both personal and biological, and a natural part of human nature. The process of naturalization is not a threat to neither personhood nor ethics if the natural relation of subjectivity and ethics is taken into account. The fact that human being live by values

that depend upon the subjective nature of those beings must give the notion of subjectivity a legitimate claim on being part of the process of naturalizing human personhood. On the other hand, subjectivity and ethics cannot exclude themselves from this process by asserting the 'unscientific' nature of the human person. My claim that human values are natural implies that the empirical sciences can clarify important aspects of the embodied nature of human personhood.

The relation between subjectivity and ethics emphasizes the necessity of a theory that does not exclude, but emphasize a patient interdisciplinary approach to human nature that accepts the many layers of our conception of what is real and natural.

References

The abbreviations used for Ricoeur's works

Books

CR: *The Course of Recognition*. Trans. D. Pellauer. Cambridge, Mass.: Havard University Press 2005; *Parcours de la reconnaissance. Trois études*. Paris: Éditions Stock 2004.

FM: *Fallible Man: Philosophy of Will*. Trans. C. A. Kelbley. New York: Fordham University Press 1987; *Philosophie de la volonté II. L'homme faillible*. Paris: Aubier 1960.

FN: *Freedom and Nature. The Voluntary and the Involuntary*. Trans. E. V. Kohák. Evanston, III: Northwestern University Press 1966; *Philosophie de la volonté I. Le volontaire et l'involontaire*. Paris: Aubier 1950.

FP: *Freud and Philosophy: Essay on Interpretation*. Trans. D. Savage. New Haven: Yale University Press 1977; *De Interpretation. Essay sur Freud*. Paris: Édition du Seuil 1965.

OSA: *Oneself as Another*. Trans. K. Blamey. Chicago: University of Chicago Press 1992; *Soi-même comme un autre*. Paris: Édition du Seuil 1990.

SE: *The Symbolism of Evil*. Trans. E. Buchanan. New York: Beacon Press 1969; *Philosophie de la volonté II. Symbolique du mal*. Paris: Aubier 1960.

TN3: *Time and Narrative, III*. Trans. K. Blamey and D. Pellauer. Chicago: University of Chicago Press 1988; *Temps et Récit, t. 3*. Paris: Édition du Seuil 1991 (1th ed. 1985).

WT: (with Jean-Pierre Changeux) *What Makes Us Think?* Trans. M. B. DeBevoise. Princeton: Princeton University Press 2000; *Ce qui nous fait penser. La nature et la regle*. Paris: Éditions Odile Jacob 2000 (1th ed. 1998).

Articles

APH: Husserl's Ideas II: Analyses and Problems. In P. Ricoeur: *Husserl. An Analysis of His Phenomenology*. Trans. E. G. Ballard and L. E. Embree (35-81). Evanton, Illinois: Northwestern University Press 2007; Analyses et problèmes des *Ideen* II de Husserl (1952). In P. Ricoeur: *A l'école de la Phénoménologie* (87-140). Paris: Vrin 1986.

EH: Exisitence and Hermeneutics. Trans. K. McLaughlin (3-24). In P. Ricoeur: *The Conflict of Interpretations*, ed. Don Ihde. New York: Continum 2004; Existence et Hérméneutique (1965). In Paul Ricoeur: *Conflit des l'interprétations* (7-28). Paris: Édition du Seuil 1969.

IDA: L'imagination dans discours et dans l'action. In P. Ricoeur: *Du texte à l'action. Essais d'herméneutique II* (237-262). Paris: Édition du Seuil 1986.

KH: Kant and Husserl. In P. Ricoeur: *Husserl. An Analysis of His Phenomenology*. Trans. E. G. Ballard and L. E. Embree (175-201). Evanton, Illinois: Northwestern University Press 2007; Kant et Husserl (1955). In P. Ricoeur: *A l'école de la Phénoménologie* (227-250). Paris: Vrin 1986.

LS: Le Sentiment (1959). In P. Ricoeur: *A l'école de la Phénoménologie* (251-265). Paris: Vrin 1986.

MT: The model of the text: meaningful action considered as a text. In Paul Ricoeur: *Hermeneutics and the human sciences*, ed. J. B. Thompson (197-221). New York: Cambridge University Press 1981; Le modèle du texte: l'action sensée considérée comme un texte (1971). In P. Ricoeur: *Du texte à l'action. Essais d'herméneutique II* (205-236). Paris: Édition du Seuil 1986.

NPH: Narrativité, Phénoménologie et Herméneutique (1989). In André Jacob (ed.): *Encyclopédie Philosophie Universelle I: L'univers philosophique* (63-71). Paris: PUF 1989.

NA: Negativity and Primary Affirmation. In P. Ricoeur: *History and Truth* (305-328). Trans. Ch. A. Kelbley. Evanton, Illinois: *Northwestern University Press 2007; Négativité et affirmation originaire (1956).* In P. Ricoeur: *Histoire* et Verité (378-405). Paris: Édition du Seuil 1967.

PD: The Problem of Double Meaning as Hermeneutic Problem and as Semantic Problem. Trans. K. McLaughlin. In P. Ricoeur: *The Conflict of Interpretations*, ed. D. Ihde (61-76). New York: Continum 2004; Le problème du double sens comme problème herméneutique et comme problème sémantique (1966). In P. Ricoeur: *Conflit des l'interprétations* (64-79). Paris: Édition du Seuil 1969.

PH: Phenomenology and Hermeneutics. Trans. J. B. Thompson. In P. Ricoeur: *Hermeneutics and the human sciences*, ed. J. B. Thompson (101-128). New York: Cambridge University Press 1981; Phénomenologie et hermémenutique (1975). In P. Ricoeur: *Du texte à l'action. Essais d'hermeneutique II* (43-81). Paris: Édition du Seuil 1986.

PW: Methods and Tasks of a Phenomenology of the Will. In *Husserl. An Analysis of His Phenomenology*. Trans. E. G. Ballard and L. E. Embree (213-233). Evanton, Illinois: Northwestern University Press 2007; Méthode et tâche d'une phénoménologie de la volonté (1951). In *Paul Ricœur: A l'école de la Phénoménologie* (59-86). Paris: Vrin 1986.

QS: The Question of the Subject: The Challenge of Semiology. Trans. K. McLaughlin. In *The Conflict of Interpretations* (232-62). Ed. D. Ihde. New York: Continum 2004; La question du sujet: le défi de la sémiologie. In Paul Ricoeur: *Conflit des l'interprétations* (pp. 233-62). Paris: Édition du Seuil 1969.

SP: Sur la Phénoménologie (1953). In P. Ricoeur: *A l'école de la Phénoménologie* (266-283). Paris: Vrin 1986.

SR: Sympathie et respect (1954). In P. Ricoeur: *A l'école de la Phénoménologie* (141-159). Paris: Vrin 1986.

TFA: True and False Anguish. In *History and Truth*. Trans. Ch. A. Kelbley (287-304). Evanton, Illinois: Northwestern University Press 2007; Vrai et fausse angoisse (1953). In P. Ricoeur: *Histoire et Verité* (357-377). Paris: Édition du Seuil 1967.

Other references

Agamben, G. (1998/1995). *Homo Sacer. Sovereign Power and Bare Life*. Trans. D. Heller-Roazen. Stanford, CA: Stanford University Press; *Homo sacer. Il potere sovrano e la nuda vita*. Torino: Einaudi.

Agamben, G. (2004/2002). *The Open. Man and Animal*. Trans. Kevin Attell. Stanford, CA: Stanford University Press; L'aperto. *L'uomo e l'animale*. Torino: Bollati Boringhieri.

Agamben, G. (2007). *Il Regno e la Gloria. Per una genealogia teologica dell'economia e del governo. Homo sacer, II, 2*. Vicenza: Neri Pozza Editore.

Aristotle (1984). Nichomachean Ethics. Trans. W. D. Ross (rev. J. O. Urmson). In J. Barnes (ed.): *The Complete Works of Aristotle. The Revised Oxford Translation*. Princeton: Princeton University Press.

Barry, B. (2005). *Why Social Justice Matters*. Cambridge: Polity Press.

Bear, Connors and Paradiso (2006). Neuroscience: *Exploring The Brain. Third Edition*. Baltimore: Lippincott Williams & Wilkins.

Beedie, C. J., Terry P. C, Lane A. M. (2005). Distinctions between emotion and mood. *Cognition and Emotion* vol. 19**,** no. 6: 847-878.

Berlucchi, G. (2007). The myth of the clonable human brain. In Della Sala, S. (ed.): *Tall Tales about the Mind and Brain. Separating fact from fiction* (336-347). Oxford: Oxford University Press.

Blackburn, S. (2006). Must We Weep for Sentimentalism? In J. Dreier (ed.): *Contemporary Debates in Moral Theory* (144-159). Oxford: Blackwell Publishing.

Blackburn, S. (2004). *Lust*. Oxford: Oxford University Press.

Blackburn, S. (1999). *Think. A compelling introduction to philosophy*. Oxford: Oxford University Press.

Blackburn, S. (1998). *Ruling Passion. A Theory of Practical Reasoning*. Oxford: Oxford University Press.

Blackburn, S. (1995). Practical Tortoise Raising. *Mind,* New Series, vol. 104, no. 416: 695-711.

Blackburn, S. (1988). How To Be an Ethical Anti-Realist. In S. Blackburn: *Essays in Quasi-Realism*. (166-181) Oxford: Oxford University Press 1993.

Blackburn, S. (1987). Morals and Modals. In S. Blackburn: *Essays in Quasi-Realism*. (52-74) Oxford: Oxford University Press 1993.

Blackburn, S. (1985). Errors and the Phenomenology of Value. In S. Blackburn: *Essays in Quasi-Realism*. (149-165) Oxford: Oxford University Press 1993.

Blackburn, S. (1984). *Spreading the Word*. *Groundings in the Philosophy of Language*. Oxford: Oxford University Press.

Brezzi, F. (2006). *Introduzione a Ricoeur*. Roma-Bari: Editori Laterza.

Cannon, W. B. (1927). The James-Lange Theory of Emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, vol. 39: 106-124.

Casebeer, W. D. (2003a). *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition*. Cambridge, Mass.: The MIT Press.

Casebeer, W. D. (2003b). Moral cognition and its neural constituents. *Nature Reviews Neuroscience*, vol. 4: 840-047.

Casebeer, W. D (2004). Knowing Evil When You See It: Uses for the Rhetoric of Evil in International Relations. *International Relations*, vol. 18, no. 4: 441-451.

Churchland, P. (2006): Moral decision-making and the brain. In J. Illes (ed.): *Neuroethics: Defining the issues in theory, practice, and policy* (3-16). New York: Oxford University Press.

Crane, T. (1998). Intentionality as the Mark of the Mental. In A. O'Hear (ed.): *Current Issues in Philosophy of Mind* (229-51). Cambridge: Cambridge University Press.

Christiansen, Lars (1980). *Naturfilosofi. Et essay om de tavse humaniora*. Købehavn: G.E.C. Gads forlag.

Damasio, A. R. (2004). Emotions and Feelings: A Neurobiological Perspective. In A. S. R. Manstead, N. Frijda, A. Fischer (eds.): *Feelings and Emotions. The Amsterdam Symposium* (49-57). Cambridge: Cambridge University Press.

Damasio, A. R. (2003). *Looking for Spinoza. Joy, Sorrow, and the Feeling Brain*. New York: Harcourt.

Damasio, A. R. (2000). A Second Chance for Emotion. In R. Lane and L. Nadel (eds.): *Cognitive Neuroscience of Emotions*. Oxford: Oxford University Press.

Damasio, A. R. (1999). *The Feeling of What happens. Body and Emotion in the Making of Consciousness*. New York: Harcourt.

Damasio, A. (1994). *Descartes' Error*. *Emotion, Reason, and the Human Brain*. New York: Putnam.

Dennett, D. C. (1991). Lovely and Suspect Qualities. *Philosophical Issues*, vol. 1, Consciousness: 37-43.

Dennett, D. C. (2003). Freedom Evolves. New York: Putnam.

de Sousa, R. (1987). *The Rationality of Emotions*. Cambridge, Mass.: MIT Press.

de Sousa, R. (2005). The Art of the Possible in Life and Literature. In M.E. Reicher and J.C. Marek (eds.): *Experience and Analysis* (347-356). Kirchberg am Wechsel, Wien: öbv&hpt.

de Sousa, R. (2007a). *Why Think? Evolution and the Rational Mind*. Oxford: Oxford University Press.

de Sousa, R. (2007b). Truth, Authenticity, and Rationality. *dialectica*, vol. 61, no. 3: 323-345.

Edelman, G. M. (2003). Naturalizing Consciousness: A Theoretical Framework. *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 100, no. 9: 5520-5524.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, vol. 6, no. 3-4: 169-200.

Ekman, P. (2003). *Emotions Revealed: Understanding Faces and Feelings*. London: Weidenfeld & Nicolson.

Farah, M. J. and Heberlein, A. S. (2007). Personhood and Neuroscience: Naturalizing or Nihilating? *American Journal of Bioethics*, vol. 7, no. 1: 37-48.

Farah, M. J. (2005). Neuroethics: the practical and the philosophical. *TRENDS in Cognitive Sciences*, vol. 9, no. 1: 34-40.

Faulkner, W. (1996/1935). *As I Lay Dying*. London: Vintage (Random House).

Fink, H. (2006). Three Sorts of Naturalism. *European Journal of Philosophy*, vol.14, no. 2: 202-221.

Franfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *Journal of Philosophy* vol. 68, no. 1: 5-20.

Fuchs, T. (2000). *Leib, Raum, Person. Entwurf einer phänomenologischen Anthropologie*. Stuttgart: Klett-Cotta.

Fuch, T. (2002a). Melancholie als Desynchronisierung. Ein Beitrag zur Psychopathologie der intersubjektiven Zeit. In T. Fuchs, *Zeit-Diagnosen. Philosophisch-Psychiatrische Essays* (111-134). Kusterdingen: Die Graue Edition.

Fuchs, T. (2002b). Maske, Selbst, Selbsentfremdung. Zur Anthropologie und Psychopathologie der Person. In T. Fuchs, *Zeit-Diagnosen. Philosophisch-Psychiatrische Essays* (135-163). Kusterdingen: Die Graue Edition.

Gallagher, S. and Marcel A. J. (1999). The Self in Contextualized Action. *Journal of Consciousness Studies* vol. 6, no 4: 4-30.

Gallagher, S. and Zahavi, D. (2008). *The phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London: Routledge.

Gallese, V. (2003). The Roots of Empathy: The Shared Manifold Hypothesis and the Neural Basis of Intentionality. *Psychopathology*, vol. 36: 171-180.

Gallese, V. (2006). Intentional Attunement: A neurophysiological perspective on social cognition and its disruption in autism. *Brain Research*, vol. 1079: 15-24.

Gallese, V. (2007). Embodied simulation: from mirror neuron systems to interpersonal relations. *Novertis Foundation Symposium*, vol. 278: 3-19.

Gallese, V. and Umiltà, M. A. (2006). Cognitive Continuity in Primate Social Cognition. *Biological Theory*, no. 1, vol. 1: 25-30.

Gazzaniga, M. S. (2006a). *The Ethical Brain. The Science of Our Moral Dilemmas*. New York: Harper Collins Publishers.

Gazzaniga, M. S. (2006b). Facts, fictions, and the future of neuroethics. In J. Illes (ed.): *Neuroethics: Defining the issues in theory, practice, and policy* (141-148). New York: Oxford University Press.

Goldie, P. (2000a). *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

Goldie, P. (2000b). Explaining Expressions of Emotion. *Mind*, New Series, vol. 109, no. 433: 25-38.

Goldie, P. (2002a). Emotion, Personality and Simulation. In P. Goldie (ed.): *Understanding Emotions: Mind and Morals* (97-109). Aldershot: Ashgate Publishing.

Goldie, P. (2002b). Emotions, feelings and intentionality. *Phenomenology and the Cognitive Sciences*, vol. 1: 235-254.

Goldie, P. (2003a). One's Remembered Past: Narrative Thinking, Emotion, and the External Perspective. *Philosophical Papers*, vol. 32, no. 3: 301-319.

Goldie, P. (2003b). Narrative, Emotion, and Perspective. In M. Kieran and D. Lopes (eds.): *Imagination and the Arts* (54-68). London: Routledge.

Goldie, P. (2003c). Narrative and perspective: Values and appropriate Emotions. In A. Hatzimoysis (ed.): *Philosophy and the Emotions* (201-220), Royale Institute of Philosophy Supplements Series. Cambridge: Cambridge University Press.

Goldie, P. (2004). *On Personality*. London: Routledge.

Goldie, P. (2005). Imagination and the Distorting Power of Emotion. *Journal of Consciousness Studies*, vol. 12, no. 8-10: 127-39.

Goldie, P. (2006). *Review of Gut Reactions: A Perceptual Theory of Emotions by J. Prinz.* Mind, vol. 115: 453-457.

Goldie, P. (2007). There are Reasons and Reasons. In D. Hutto and M. J. Ratcliffe: *Folk Psychology Reassessed* (103-114). Dordrecht Netherlands: Kluwer Academic Press.

Goldman, A.I. and Sripada, C. S. (2005). Simulationist models of faced-based emotion recognition. *Cognition*, vol. 94: 193-213.

Greene, J. (2003). From neural "is" to moral "ought": what are the moral implications of neuroscientific moral psychology? *Nature Reviews Neuroscience*, vol. 4: 847-850.

Greisch, J. (2000). *Le Cogito herméneutique. L'herméneutique philosophique et l'héritage cartésien*. Paris: Libraire philosophique J. Vrin.

Greisch, J. (2001). *Paul Ricoeur. L'itinérance du sens*. Grenoble: J. Milton.

Grøn, A. (2004a). The Embodied Self: Reformulating the Existential Difference in Kierkegaard. *Journal of Consciousness Studies*, vol. 11, no. 10-11: 26-43.

Grøn, A. (2004b). Self and Identity. In D. Zahavi, T.Grünbaum, J. Parnas: *The Structure and Development of Self-Consciousness*. *Interdisciplinary perspectives* (123-156). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Heidegger, M. (1990/1998). *Kant and the Problem of Metaphysics*. Trans. R. Taft. Indiana: Indiana University Press; *Kant und das Problem der Metaphysik*. Frankfurt am Main: Vittorio Klostermann.

Heidegger, M. (1996/1993). *Being and Time*. Trans. J. Stambaugh. Albany, NY: State University Of New York; *Sein und Zeit*. Tübingen: Max Niemeyer Verlag.

Hume, D. (1751/1975). *Enquiries concerning Human Understanding and Concerning the Principles of Morals*. Oxford: Clarendon Press.

Hume, D. (1739/1978). *A Treatise of Human Nature*. Oxford: Clarendon Press.

Jacobs, J. (2002). *Dimensions of Moral Theory. An Introduction to Metaethics and Moral Psychology*. Oxford: Blackwell.

James, W. (1884). What is an Emotion? *Mind*, vol. 9, no. 34: 188-205.

Jervolino, D. (1993). *Il cogito e l'ermeneutica. La questione del soggetto in Ricoeur*. Genova: Marietti.

Jervolino, D. (1995). *L'amore difficile*. Roma: Studium.

Kant, I. (1996/1914). *The Metaphysics of Morals*. Trans. M. Gregor. Cambridge: Cambridge University Press; *Die Metaphysik der Sitten*. Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 6. Berlin: W. de Gruyter.

Kant, I. (1996/1914). *The Metaphysics of Morals*. Trans. M. Gregor. Cambridge: Cambridge University Press; *Die Metaphysik der Sitten*. Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 6. Berlin: W. de Gruyter.

Kant, I. (1997/1904). *Critique of Pure Reason*. Trans. P Guyer. Cambridge: Cambridge University Press; *Kritik der reinen Vernunft*. Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 3. Berlin: W. de Gruyter.

Kant, I. (1998a/1903). *Groundwork of the Metaphysics of Morals*. Trans. P Guyer. Cambridge: Cambridge University Press; *Grundlegung zur Metaphysik der Sitten* . Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 4. Berlin: W. de Gruyter.

Kant, I. (1998b/1908). *Critique of Practical Reason*. Trans. M. Gregor. Cambridge: Cambridge University Press; *Kritik der praktischen Vernunft*. Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 5. Berlin: W. de Gruyter.

Kant, I. (2006/1917). *Anthropology from a Pragmatic Point of View*. Trans. R. B. Louden. Cambridge University Press; *Anthropologie in pragmatischer Hinsicht*. Immanuel Kants Schriften. Ausgabe der könglich Preussischen Akademie der Wissenschaften, band 7. Berlin: W. de Gruyter.

Kenny, A. (1963). *Action, Emotion and Will*. London: Routledge.

Kennedy, D. (2002). Are There Things We'd Rather Not Know? In Marcus, S. J. (ed.): *Neuroethics: Mapping the Field* (196-206). New York: Dana Foundation.

Kierkegaard, S. (1983/1920). *The Sickness unto Death*. Trans. H.V. Hong and E. H. Hong. Princeton, NJ: Princeton University Press; *Sygdommen til Døden*. Søren Kierkegaards Samlede Værker, Anden Udgave, bind 11. København: Gyldendalske Boghandel.

Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.

LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional life*. New York: Simon and Schuster.

Legrand, D. (2007). Naturalizing the Acting Self: Subjective vs. Anonymous Agency. *Philosophical Psychology*, vol. 20, no. 4: 457-478.

Levy, N. (2007). *Neuroethics*. *Challenges for the 21st Century*. Cambridge: Cambridge University Press.

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. Harmondsworth, Middx.: Penguin Books.

Marcus, S. J. (ed.) (2002). *Neuroethics: Mapping the Field*. New York: Dana Foundation.

McDowell, J. (1997). Projection and Truth in Ethics. In J. McDowell: *Mind, Value, and Reality* (151-166). Cambridge, Mass.: Harvard University Press.

Metzinger, Th. (2005). Unterwegs zu einem neuen Menschenbild. *Gehirn&Geist*, no. 11.

Mongin, O. (1994). *Paul Ricoeur*. Paris: Édition du Seuil.

Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.

Murdoch, I. (1967): The Sovereignty of Good over Other Concepts. In I. Murdoch: *The Sovereignty of Good* (77-104). London: Routledge 1970.

Nagel, T. (1986). *The View from Nowhere*. New York: Oxford University Press.

Nietzsche, F. (1986/1954a). *Human, all too human*. Trans. R.J. Hollingdale. Cambridge: Cambridge University Press; *Menschliches, Allzumenschliches*. Werke in drei Bänden. Hrsg. von Karl Schlechta, Band 1. Münschen: Carl Hanser Verlag.

Nietzsche, F. (2007/1954b). *The Dawn*. Trans. J. M. Kennedy. New York: Dover Publications; *Morgenröte*. Werke in drei Bänden. Hrsg. von Karl Schlechta, Band 1. Münschen: Carl Hanser Verlag.

Nussbaum, M. (2001). *Upheavals of Thought*. Cambridge: Cambridge University Press.

Overgaard, S. (2007). *Wittgenstein and Other Minds. Rethinking subjectivity and intersubjectivity with Wittgenstein, Levinas, and Husserl*. London: Routledge.

Orwell, G. (1989/1949): *Nineteen Eighty-Four*. London: Penguin Books.

Panksepp, J. (1991). Affective Neuroscience: A conceptual framework for the neurobiological study of emotions. In K. Strongman (ed.): *International Reviews of Emotional Research* (59-99). Chichester: Wiley.

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford: Oxford University Press.

Panksepp, J. (2001). On the Subcortical Sources of Basic Human Emotions and the Primacy of Emotional-Affective Action-Perception Processes in Human Consciousness. *Evolution and Cognition*, vol. 7, No. 2: 134-140.

Panksepp, J. (2003). At the interface of the affective, behavioural, and cognitive neuroscience: Decoding the emotional feelings of the brain. *Brain and Cognition*, vol. 52: 4-14.

Panksepp, J. (2004a). Affective Consciousness and the Origins of Human Mind: A Critical Role of Brain Research on Animal Emotions. *Impuls*, no. 3: 47-60.

Panksepp, J. (2004b). Basic Affects and the Instinctual Systems of the Brain. In A. S. R. Manstead, N. Frijda, A. Fischer (eds.): *Feelings and Emotions. The Amsterdam Symposium*, eds. (175-193). Cambridge: Cambridge University Press.

Panksepp, J. (2005a). Toward a science of ultimate concern (Guest Editorial). *Consciousness and Cognition*, vol. 14: 22-29.

Panksepp, J. (2005b). Affective Consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition*, vol. 14: 30-80.

Panksepp, J. (2005c). On the Embodied Neural Nature of Core Emotional Affects. *Journal of Consciousness Studies*, vol. 12, no. 8-10: 158-84.

Panksepp, J. (2006). Emotional endophenotypes in evolutionary psychiatry. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, vol. 30: 774-784.

Panksepp, J. (2007). Neuroevolutionary sources of laughter and social joy: Modelling primal human laughter in laboratory rats. *Behavioural Brain Research*, vol. 182: 231-244.

Phelps, E. A. (2004). The Human Amygdala and Awareness: Interactions Between Emotion and Cognition. In M. S. Gazzaniga (ed.): *The Cognitive Neurosciences III* (1005-1015). Cambridge: The MIT Press.

Pinker, S. (1999). *How The Mind Works*. New York: Penguin Books.

Pinker, S. (2003). *The Blank Slate. The Modern Denial of Human Nature*. New York: Penguin Books.

Preuss, T. M. (2004). What is it like to be a Human? In M. S. Gazzaniga (ed.): *The Cognitive Neurosciences III* (5-19). Cambridge: The MIT Press.

Prinz, J. J. (2002). *Furnishing the Mind: Concepts and their Perceptual basis*. Cambridge, MA: MIT Press.

Prinz, J. J. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. New York: Oxford University Press.

Prinz, J. J. (2005). Are Emotions Feelings?. *Journal of Consciousness Studies* **12**, no. 8-10:9-25.

Prinz, J.J. (2007). *The Emotional Construction of Morals*. New York: Oxford University Press.

Ratcliffe, M. (2005). The Feeling of Being. *Journal of Consciousness Studies* **12**, no. 8-10: 43-60.

Reich, R. B. (2007). *Supercapitalism. The Transformation of Business, Democracy, and Everyday Life*. New York: Alfred A. Knoff.

Rizzolatti, G., Fogassi, L. and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews, vol. 2: 661-669.

Rizzolatti, G. and Sinigaglia, C. (2007). *Mirrors in the Brain. How our Minds share Action, Emotions, and Experience*. Trans. F. Anderson. Oxford: Oxford University Press.

Rochat, P. (2007). Intentional action arises from early reciprocal exchanges. *Acta Psychologica*, vol. 124, no. 1: 8-25.

Rosfort, R. and Stanghellini, G. (forthcoming). The Person in between Moods and Affects. *Philosophy, Psychiatry, and Psychology*.

Roskies, A. (2002). Neuroethics for the New Millinium. *Neuron*, Vol. 35: 21-23.

Roskies, A. (2006): A case study of neuroethics: the nature of moral judgment. In J. Illes (ed.): *Neuroethics: Defining the issues in theory, practice, and policy* (17-32). New York: Oxford University Press.

Rudrauf, D. and Damasio, A. (2005). A Conjecture Regarding the Biological Mechanism of Subjectivity and Feeling. *Journal of Consciousness Studies* vol.12**,** no. 8-10: 26-42.

Russell, B. (1960). *An Outline of Philosophy*. Cleveland: The World Publishing Company.

Russell, J. A. (2005). Emotion in Human Consciousness Is Build on Core Affect. In *Journal of Consciousness Studies* vol. 12, no. 8-10:26-42.

Russell, J. A. (2003). Core affect and the Psychological Construction of Emotion. *Psychological Review* vol. 110, no**.** 1: 145-172.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.

Safire, W. (2002): Visions for a New Field of "Neuroethics". In Marcus, S. J. (ed.): *Neuroethics: Mapping the Field* (3-9). New York: Dana Foundation.

Sartre, Jean-Paul (1971/1939). *Sketch for a Theory of the Emotions.* Trans. Philip Mairet. London: Methuen & Co Ltd; *Esquisse d'une théorie des émotions*. Paris: Hermann.

Scheler, M. (1973/1912). *Wesen und Formen der Sympathie*. Max Scheler. Gesammelte Werke. Band 7. Bern: Francke Verlag.

Scheler, M. (1966/1916). *Der Formalismus in der Ethik und die Materiale Wertethik*. Max Scheler. Gesammelte Werke. Band 2. Bern: Francke Verlag.

Searle, T. (2000). *Mind, Language and Society*. London: Phoenix, Orion Books Ltd.

Siemer, M. (2005). Moods as multiple-object directed and as objectless affective states: An examination of the dispostional theory of moods. *Cognition and Emotion* vol. 19, no. 6: 815-845.

Smith, Q. (1986). *The Felt Meanings of the World. A Metaphysics of Feeling*. Indiana: Purdue University Press.

Solomon, R. C. (1976). *The Passions*. New York: Anchor Press.

Solomon, R. C. (1977). The Logic of Emotion. *Noûs*, vol. 11, no. 1: 41-49.

Solomon, R. C. (2001). Thoughts and Feelings: What is a "Cognitive Theory" of the Emotions, and Does It Neglect Affectivity? In R. Solomon: Not Passions Slave (178-194). Oxford: Oxford university Press 2003.

Solomon, R. C. (2007). *True to Our Feelings. What Our Emotions are really telling Us*. New York: Oxford University Press.

Simpson, G. G. (1966). The Biological Nature of Man. *Science*, no. 152: 472-8.

Spinoza, B. (1955). *The Ethics*. Chief works vol. 2. Trans. R.H.M. Elwes. New York: Dover Publications.

Stanghellini, G. (1997). *Antropologia della vulnerabilità*. Milano: Feltrinelli.

Stanghellini, G. (2004). *Disembodied Spirits and Deanimated Bodies: The Psychopathology of Common Sense*. Oxford: Oxford University Press.

Stanghellini, G. (2006). *Pscicopatologia del senso comune*. Milano: Raffaello Cortina Editore.

Stanghellini, G. and Ballerini, A. (1992). *Ossessione e rivelazione. Riflessione sui rapporti tra ossessività e delirio*. Torino: Bollati Boringhieri.

Stocker, M. (1976). The Schizophrenia of Moral Ethical Theories. *The Journal of Philosophy*, vol. 73, no. 14: 453-466.

Stocker, M. (1983). Psychic Feelings: Their Importance and Irreducibility. *Australasian Journal of Philosophy*, vol. 61, no. 1: 5-26.

Stocker, M. (1996), with Hegeman, E. *Valuing Emotions*. Cambridge: Cambridge University Press.

Strasser, S. (1977/1956). *Phenomenology of Feeling: An essay on the Phenomena of the Heart*. Trans. R.E. Wood. Pittsburgh: Duquesne University Press; *Das Gemüt: Grundgedanken zu einer Phänomenologische Philosophie und Theorie des Menschlichen Gefühlsleben*. Freiburg: Verlag Herder.

Strawson, G. (1999). The Self and the SESMET. *Journal of Consciousness Studies*, vol. 6, no. 4: 99-135.

Strawson, G. (2004). Against Narrativity. *Ratio (new series)*, vol. 16, no. 4: 428-452.

Strawson, P. F. (1959). *Individuals. An Essays in Descriptive Metaphysics*. London: Methuen & Co. Ltd.

Stawson, P. F. (1962). Freedom and Resentment. In G. Watson (ed.): *Free Will* (59-80). Oxford: Oxford University Press 1982.

Strawson, P. F. (1966). *The Bounds of Sense. An Essay on Kant's Critique of Pure Reason*. London: Methuen & Co. Ltd.

Taylor, C. (1977a). Self-interpreting animals. In C. Taylor: *Human Agency and Language. Philosophical Papers 1* (45-76). Cambridge: Cambridge University Press 1985.

Taylor, C. (1977b). What is human agency? In C. Taylor: *Human Agency and Language. Philosophical Papers 1* (15-44). Cambridge: Cambridge University Press 1985.

Taylor, C. (1981). The concept of a person. In C. Taylor: *Human Agency and Language. Philosophical Papers 1* (97-114). Cambridge: Cambridge University Press 1985.

Taylor, C. (1988). Reply to Sousa and Davis. *Canadian Journal of Philosophy*, vol. 18, no. 3: 449-458.

Taylor, C. (1989). *Sources of the Self. The Making of Modern Identity*. Cambridge: Cambridge University.

Teichert, D. (2004). Narrative, Identity and the Self. *Journal of Consciousness Studies*, vol. 11, no. 10-11: 175-191.

Watson, J. D. and Crick F. H. (1953): Molecular Structure of Nucleic Acids. *Nature*, vol. 171, no. 4356: 737-738.

Wollheim, R. (1984). *Thread of Life*. Cambridge, Mass.: Havard University Press.

Wollheim, R. (1999). *On the Emotions*. New Haven: Yale University Press.

Zahavi, D. (2005). *Subjectivity and Selfhood. Investigation the First-Person Perspective*. Cambridge, Mass.: The MIT Press.

Zahavi, D. (2007). Killing the straw man: Dennett and phenomenology. *Phenomenology And The Cognitive Sciences*, vol. 6, no. 1-2: 21-43.

Zoloth, L. (2006). Being in the world: Neuroscience and the ethical agent. In J. Illes (ed.): *Neuroethics: Defining the issues in theory, practice, and policy* (61-73). New York: Oxford University Press.

Dissertation Abstract

Subjectivity and Ethics. Ricoeur and the Question of Naturalizing Personhood

This dissertation argues for the irreducible relation between subjectivity and ethics. To deny subjectivity a place in ethics is to misunderstand the nature of ethical theory and to exclude ethics from a theory of subjectivity is to disregard a critical feature of subjective experience, namely the question of normativity. Human experience of the world is normative. We do not experience the world as a value-free space of indifferent events and movements, but as meaningful situations and actions that inscribe themselves in a common world of interest, care, and value for the subject. The human subject is characterized by what it cares about, and subjective experience is primarily structured in accordance with the values that are expressed in the basic notion of care. I therefore begin the dissertation by performing a systematic reformulation of Ricoeur's theory of subjective in order to provide a conceptual framework for my further analyses of subjectivity and ethics.

Ethical experience is critical in the structure of subjective experience, since subjective values, albeit they origin in the self, are never restricted to the sole being of the individual subject. Subjective experience is constituted by selfhood and alterity. I do not experience the world or myself as creations of my will. Thus, my being is affected by the presence of alterity in the sense that selfhood, world, and the other are all active parts in the constitution of the values by which I lead my life. I cannot neglect this alterity in the constitution of my being, since my experience is continually affected by that which is not me (my birth, my fragmentary perspective, my physical body, time, space, and other persons who react to my doing and saying).

Further, subjective experience is characterized by how the subject interacts with the world that it experiences. Conscious experience is not mere registration, but the result of the relation between activity and passivity, action and suffering. Hence, I argue that subjective experience is structured as a practical space of reason according to a hierarchy of heterogeneous values generated and shaped by the interaction of the world, the other, and the self. We experience the world as a configuration of things, events, and situations that we care about.

Subjectivity and ethics is irreducibly related, because the subject is a person other than a mere physical being. Personhood is what makes the subject a being who cares about more than its own survival and wellbeing, and I argue that we cannot understand the nature of subjectivity in isolation from personhood. The fact that human subjects care about the world, the other, and themselves is what leads me to argue for an ontology of care that is capable of explaining the complexity of human nature. We need to account for both the physical and personal nature of subjectivity if we want to give a comprehensive and satisfying explanation of the existence of a human being in a world that is not human. Feelings are fundamental for understanding the structure and value-loaded nature of subjective experience and being. Therefore, I spent considerable time on analyzing both the neurophysiological and psychological dimensions of human affectivity.

I finally examine the nature of human values in order to account for the particular personal values that characterize human care. I argue for a naturalist conception of human nature and values and how naturalism is not to be considered a threat to the particular nature of human feeling, action, and personhood. I clarify my conception of naturalism by criticizing the newly born discipline called neuroethics and a naturalistic based moral relativism. The process of naturalization must respect the subjective and personal nature of human beings.

The analyses are structured in three parts:

1) A Reformulation of Ricoeur's Theory of Subjectivity (a conceptual framework: ca. 100 pages)
2) Feeling Ethical (analyses of human affectivity and its relation to ethical experience: ca. 100)
3) The Brain and Human Values (neuroethics and the Process of naturalisation: ca. 30)

Afhandlingsresumé
Subjektivitet og Etik. Ricœur og spørgsmålet om personbegrebets naturalisering

Afhandlingen argumenterer for det ikke-reducerbare forhold mellem subjektivitet og etik. At benægte subjektivitet en plads i etikken er at misforstå etisk teori og at udelukke etik fra en subjektivitetsteori er at se bort fra et vigtigt aspekt af subjektiv erfaring, nemlig spørgsmålet om det normative. Menneskelig erfaring af verden er normativ. Vi erfarer ikke verden som et værdifrit rum af begivenheder og bevægelser, men som meningsfyldte situationer og handlinger, der indskriver sig i en fælles verden af interesse, omsorg/bekymring (care) og værdi for subjektet. Det menneskelige subjekt er karakteriseret ved hvad det bekymrer sig om, og subjektiv erfaring er først og fremmest struktureret i overensstemmelse med de værdier, som er udtrykt i det fundamentale begreb *omsorg/bekymring* (care). Jeg begynder derfor afhandlingen med at foretage en systematisk reformulering af Ricœurs subjektivitetsteori, der således danner en begrebslig ramme for mine videre analyser af forholdet mellem subjektivitet og etik.

Etisk erfaring er afgørende for den subjektive erfarings struktur, fordi subjektive værdier, selvom de har deres oprindelse i selvet, er aldrig begrænset til det individuelle subjekts eksistens. Subjektiv erfaring er konstitueret af selv (selfhood) og fremmedhed (alterity). Jeg erfarer ikke verden eller mig selv som skabelser af min egen vilje. Min væren er derfor berørt (affected) af fremmedhedens nærværd i og med, at selv, verden og det andet menneske alle er aktive faktorer i konstitutionen af de værdier, hvorved jeg lever mit liv. Jeg kan ikke se bort fra denne fremmedhed i min væren, fordi min erfaring konstant er berørt af det, som ikke er mig (min fødsel, mit fragmentariske perspektiv, min fysiske krop, tid, rum og andre personer, som reagerer på mine handlinger og ord).

Dertil er subjektive erfaring is karakteriseret ved den måde, subjektet interagerer med den verden, som det erfarer. Bevidst erfaring er ikke en simpel registrering, men et resultat af forholdet mellem aktivitet og passivitet, handling og lidelse. Derfor argumenterer jeg for, at subjektiv erfaring er struktureret som et fornuftens praktiske rum (a practical space of reason) i overensstemmelse med et hierarki af heterogene værdier skabte og bearbejdede i et samspil af verden, selvet og den anden person. Vi erfarer verden som en konfiguration af ting, begivenheder og situationer, som vi har omsorg for/bekymrer os om.

Subjektivitet og etik er uløseligt forbundet, fordi subjektet er en person udover at være et fysisk væsen. At subjektet er en person betyder, at det bekymrer sig om andet og mere end dets egen overlevelse og velvære, og subjektivitet er uforståelig uden personbegrebet. Det faktum at subjektet bekymrer sig om verden, den anden og sig selv fører mig til at argumentere for en omsorgens/bekymringens ontologi (care), der er i stand til at forklare menneskets komplekse natur. Vi er nødt til at forklare både subjektets fysiske og personlige natur, hvis vi vil give en fyldestgørende forklaring på den menneskelige eksistens i en verden, der ikke er menneskelig. Følelserne er fundamentale for forståelsen af den værdiladede natur, der kendetegner subjektets erfaring og væren. Jeg bruger derfor betydelig tid på at analysere både den neurofysiologiske og psykologiske dimension af menneskelige følelser.

Jeg slutter med at undersøge de menneskelige værdiers natur for at forklare de personlige værdier, som kendetegner menneskelig bekymring. Jeg vælger en naturalistisk forståelse af menneskelige værdier og viser, hvordan naturalisering ikke er en trussel mod den særlige natur, der kendetegner menneskelig følelse, handling, og personbegreb. Jeg klargør min forståelse af naturalisme ved at kritisere henholdsvis neuroetik og en naturalistisk baseret moralsk relativisme. En naturalisering må respekterer menneskets subjektive og personlige natur. Analyserne er struktureret i tre dele:

1) En reformulering af Ricœurs subjektivitetsteori (en begrebslig ramme)
2) Føle sig etisk (analyser af menneskelige følelser og forbindelsen til etisk erfaring)
3) Hjernen og menneskelige værdier (neuroetik, moralsk relativisme og naturalisering)